

The intermediate disorder regime for a directed polymer model on a hierarchical lattice

Tom Alberts* & Jeremy Clark† & Saša Kocić‡

Abstract

We study a directed polymer model defined on a hierarchical diamond lattice, where the lattice is constructed recursively through a recipe depending on a branching number $b \in \mathbb{N}$ and a segment number $s \in \mathbb{N}$. When $b \leq s$ previous work [27] has established that the model exhibits strong disorder for all positive values of the inverse temperature β , and thus weak disorder reigns only for $\beta = 0$ (infinite temperature). Our focus is on the so-called *intermediate disorder regime* in which the inverse temperature $\beta \equiv \beta_n$ vanishes at an appropriate rate as the size n of the system grows. Our analysis requires separate treatment for the cases $b < s$ and $b = s$. In the case $b < s$ we prove that when the inverse temperature is taken to be of the form $\beta_n = \hat{\beta}(b/s)^{n/2}$ for $\hat{\beta} > 0$, the normalized partition function of the system converges weakly as $n \rightarrow \infty$ to a distribution $\mathbf{L}(\hat{\beta})$ depending continuously on the parameter $\hat{\beta}$. In the case $b = s$ we find a critical point in the behavior of the model when the inverse temperature is scaled as $\beta_n = \hat{\beta}/n$; for an explicitly computable critical value $\kappa_b > 0$ the variance of the normalized partition function converges to zero with large n when $\hat{\beta} \leq \kappa_b$ and grows without bound when $\hat{\beta} > \kappa_b$. Finally, we prove a central limit theorem for the normalized partition function when $\hat{\beta} \leq \kappa_b$.

1 Introduction

Probabilistic models of directed polymers are often constructed as modifications of directed random walk models on a particular lattice. For the particular model that we study, directed polymers in a random environment, the path probabilities are randomly perturbed by a Gibbsian reweighting that is determined by each realization of the environment, and the strength of the perturbation is controlled by an inverse temperature parameter β . Broadly speaking, the main question of interest in such models is how the presence of the random environment affects these path probabilities, and to what degree any effect can be quantified. The values of β for which the presence of the environment has no substantial effect are called the *weak disorder regime*, while those β for which the environment has a meaningful influence make up the *strong disorder regime*. The precise characterization of weak versus strong disorder is given in terms of the positivity of a martingale limit, the martingale being the normalized partition function associated with the model.

In recent years there has been substantial attention paid to determining the exact structure of the weak and strong disorder regimes. On the integer lattice \mathbb{Z}^d (which is commonly referred to as $d + 1$ -dimensional to indicate that the polymer is directed) it is known [11] that there is a critical value $\beta_c \geq 0$ such that $\beta < \beta_c$ is in the weak disorder regime and $\beta > \beta_c$ is in the strong disorder regime. When $d = 1$ or 2 it is known that the polymer is in the strong disorder regime for all finite temperatures (i.e. $\beta_c = 0$), whereas for $d \geq 3$ it is known that there is a non-trivial interval of finite temperatures that make up the weak disorder regime (i.e. $\beta_c > 0$). In the latter case there are only bounds for the value of β_c and how it depends on the dimension d and the statistical distribution of the underlying environment.

*Department of Mathematics, University of Utah: alberts@math.utah.edu

†Department of Mathematics, University of Mississippi: jeremy@olemiss.edu

‡Department of Mathematics, University of Mississippi: skocic@olemiss.edu

In contrast, for directed polymers on a self-similar tree there is an *exact* characterization [5, 26] of the value of β_c (for any choice of the statistical distribution of the environment satisfying weak moment assumptions), and much more is known about the strong and weak disorder regimes [7, 30]. The tree model is easier to analyze because of the underlying geometry of the space, namely that the tree is self-similar and that paths, once they split, can never recombine. In particular, the non-recombining feature means that the random energies assigned to different paths are essentially independent after a long time, which is very different from what happens in the lattice case.

Recently a new phenomenon has emerged in directed polymer models: the *intermediate disorder regime*. The terminology *intermediate disorder* was first introduced in [1] to indicate that it sits between weak and strong disorder, and the regime itself is accessed by scaling the inverse temperature β towards its critical value as the size of the system goes to infinity. The correct scaling produces new and interesting behavior that is fundamentally different from what happens for strong and weak disorder. At the same time, the intermediate regime often contains a full range of models that act (or are conjectured to act) as a bridge between the two classical regimes.

Typically intermediate disorder is easier to study in situations where $\beta_c = 0$, since in this case the polymer model is only a small perturbation of the underlying random walk. In $d = 1$ it was shown in [1] that the proper scaling for intermediate disorder is to replace β with $\beta n^{-1/4}$, and that under a diffusive scaling of space and time the entire random Gibbs measure converges (in law) to a random measure on continuous paths called the continuum directed random polymer. Properties of the continuum polymer were analyzed in [2], but it is important to note that the continuum polymer is actually an entire *family* of models indexed by the parameter β . It is not yet understood how or even if this family acts as a bridge between weak and strong disorder, although encouraging results in this direction were obtained in [13]. For $d = 2$ the situation for intermediate disorder is largely open, and for $d \geq 3$ it seems that little has been done for scalings around β_c . A model somewhat similar to the $d = 1$ case, the *continuum disordered pinning model* (CDPM), was considered in [9]. Again in that case $\beta_c = 0$, and the authors were able to determine the precise scaling for intermediate disorder and take a scaling limit of the random Gibbs measure. In the tree case, where $\beta_c > 0$, the proper scaling window for the intermediate disorder regime was determined in [3], although nothing could be said about infinite volume limits of the corresponding Gibbs measures.

In this paper we study intermediate disorder on the *diamond hierarchical lattice*. This particular lattice has an interesting geometrical structure that still possesses a self-similarity like the tree does, but at the same time allows paths to reintersect like they do on \mathbb{Z}^d . In this sense the diamond lattice is a natural generalization of the tree but in many important ways is actually more like \mathbb{Z}^d , and for the polymer model with site disorder (which is what we consider) this is especially true. See [28, Remark 1.1] for a convincing argument in favor of this viewpoint. The structure of the diamond lattice is controlled by two parameters $b, s \in \mathbb{N}$: the *branching number* and the *segment number*. The “lattice” is actually a sequence of graphs constructed by a recursive procedure determined by the branching and segment numbers, and the depth of the recursive determines the length of the polymer. For directed polymers on the diamond lattice it was proved in [27] that there is a β_c that strictly separates weak disorder from strong disorder, as in the $d + 1$ -dimensional case, and that $\beta_c = 0$ for $b \leq s$ while $\beta_c > 0$ for $b > s$. These results are precisely summarized in Section 2.

Our main focus is on the intermediate disorder regime in the $b < s$ and $b = s$ cases, i.e., when $\beta_c = 0$, which exhibit widely disparate behaviors. Our analysis, in a sense, extends the results of [27] to obtain a richer description of the phase diagram around the critical point. In the case $b < s$, one sees through a simple Taylor expansion of the partition function that the correct intermediate disorder scaling is to replace β with $\hat{\beta}(b/s)^{n/2}$ for a parameter $\hat{\beta} > 0$; see Section 3.3. Similarly to the $1 + 1$ -dimensional polymer and the continuum disordered pinning model, we find that there is a one-parameter family of limit laws $\mathbf{L}(\hat{\beta})$ for the normalized partition functions of the model. The limiting family depends continuously on $\hat{\beta}$ and for $b < s$ satisfies

- For $\hat{\beta} \ll 1$, $\mathbf{L}(\hat{\beta})$ is approximately a normal distribution with mean 1 and variance $\hat{\beta}^2 \frac{b(s-1)}{s-b}$.
- If $X^{(i,j)}(\hat{\beta})$ are independent random variables with distribution $\mathbf{L}(\hat{\beta})$, then

$$\mathbf{L}(\hat{\beta}\sqrt{s/b}) \stackrel{d}{=} \frac{1}{b} \sum_{i=1}^b \prod_{1 \leq j \leq s} X^{(i,j)}(\hat{\beta}).$$

As in [1, 8, 9] the limiting law appears as a *universal* object, i.e. the same limit appears for any disorder variables with exponential moments that obey proper normalization constraints, but the limit itself is not Gaussian. Furthermore these limiting laws can be used to construct a continuum diffusion process interacting with a random environment on the continuum diamond lattice, similar to what is done in [2, 9]. This is discussed in Section 4. We emphasize that in contrast to [1, 8, 9] our methods for showing the existence of these limit laws and their universality properties do not use polynomial chaos expansions; instead they come from renormalization group type ideas.

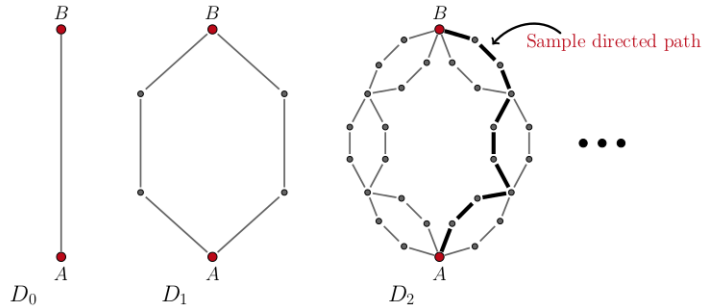
In the case $b = s$, nontrivial behavior emerges in the intermediate disorder regime by replacing β with $\hat{\beta}/n$, but in contrast to the 1 + 1-dimensional polymer and the CDPM the parameter $\hat{\beta}$ does not seem to act as any sort of bridge. In fact, in the $b = s$ case, there is a critical value κ_b such that the variance of the rescaled partition function converges to zero with large n for $\hat{\beta} \leq \kappa_b$, but also grows to infinity for $\hat{\beta} > \kappa_b$. This is discussed in Section 5.

The structure of this paper is as follows: in Section 2 we give a full description of the model and precisely state our main results. In Section 3 we describe our coordinate system for the diamond lattice and use it to derive recursive formulas for moments of the partition function. We especially point out Section 3.3, which gives some heuristic reasoning for the choice of scalings $\hat{\beta}(b/s)^{n/2}$ and $\hat{\beta}/n$ in the $b < s$ and $b = s$ cases, respectively. Section 4 contains a proof of a limit theorem for the normalized partition function in the $b < s$ case, and Section 5 contains our results on the $b = s$ case. In Sections 6 and 7 we give, respectively, very brief discussions of the $b > s$ case and the equivalent model in the $b = s$ case in which random weights are placed on the bonds of the graph rather than the sites. Finally, we relegate proofs that are particularly calculus-based to Section 8.

Acknowledgements: Tom and Saša thank Kostya Khanin for useful discussions.

2 The Model and Main Results

Diamond graphs are constructed inductively by replacing each edge in D_n by a ‘diamond’ with $b \in \mathbb{N}$ branches, each of which is split into $s \in \mathbb{N}$ segments:



The first few diamond graphs constructed with $b = 2$ and $s = 3$.

The n th diamond graph defines a set of directed paths between the root vertices A and B , which we denote by Γ_n . The inductive nature of the construction implies that D_{m+n} is equivalent to replacing each edge on D_m by a copy of D_n . We assume throughout that $b > 1$, since otherwise the graph structure is trivial.

We now place i.i.d. random variables ω_a at the set of vertices $a \in D_n$ and consider statistical mechanical quantities defined through sums over the set of directed paths, Γ_n . Since D_n is canonically embedded in D_{n+1} , we can view the random variables ω_a , $a \in \cup_{n=1}^{\infty} D_n$ as residing in a single probability space. We assume that the random variables have mean zero, variance one, and finite exponential moments

$$\lambda(\beta) := \log \mathbb{E}[e^{\beta \omega_a}] < \infty$$

for sufficiently small $\beta > 0$. The partition function $Z_n(\beta)$ is the random positive quantity defined by

$$Z_n(\beta) := \frac{1}{|\Gamma_n|} \sum_{p \in \Gamma_n} e^{\beta H_n^\omega(p)}, \quad \text{where} \quad H_n^\omega(p) := \sum_{a \in p} \omega_a \quad \text{and} \quad |\Gamma_n| = b^{\frac{s^n - 1}{s - 1}}.$$

The partition function normalizes random probability measures $\mu_{\beta,n}^{(\omega)}$ on the path space $p \in \Gamma_n$:

$$\mu_{\beta,n}^{(\omega)}(p) = \frac{e^{\beta H_n^\omega(p)}}{Z_n(\beta)}.$$

Disorder for directed polymer models is commonly interpreted through a closely related quantity, $W_n(\beta)$, defined by

$$W_n(\beta) := \frac{Z_n(\beta)}{\mathbb{E}[Z_n(\beta)]}, \quad \text{which can be written as} \quad \frac{1}{|\Gamma_n|} \sum_{p \in \Gamma_n} \prod_{a \in p} E_a(\beta) \quad (2.1)$$

for $E_a(\beta) := e^{\beta \omega_a - \lambda(\beta)}$. The inductive nature of the construction of the diamond graphs D_n suggests that there should be a recursive relation between the distributions of $W_n(\beta)$ and $W_{n+1}(\beta)$; if $W_n^{(i,j)}(\beta)$ are independent copies of $W_n(\beta)$ and $E^{(i,j)}(\beta)$ are independent copies of $E_a(\beta)$, then there is the following equality in distribution

$$W_{n+1}(\beta) \stackrel{d}{=} \frac{1}{b} \sum_{i=1}^b \prod_{1 \leq j \leq s} W_n^{(i,j)}(\beta) \prod_{1 \leq j \leq s-1} E^{(i,j)}(\beta). \quad (2.2)$$

Moreover, the sequence of random variables W_n forms a mean one martingale with respect to the filtration $\mathcal{F}_n := \sigma\{\omega_a \text{ for } a \in D_m, m \leq n\}$. The martingale convergence theorem guarantees the almost sure existence of a nonnegative limit as $n \rightarrow \infty$:

$$W_\infty(\beta) = \lim_{n \rightarrow \infty} W_n(\beta).$$

The literature [6, 32, 27] on random polymer models commonly interprets those β for which $W_\infty(\beta) > 0$ as being in the *weak disorder* regime and those β for which $W_\infty(\beta) = 0$ as being in the *strong disorder* regime.

In Lemma 2.4 and Proposition 2.7 of [27], Lacoïn and Moreno establish the following results for the diamond hierarchical lattice (amongst others):

- i) For any fixed b , s , and β , the random variable $W_\infty(\beta)$ satisfies the zero-one law

$$\mathbb{P}[W_\infty(\beta) > 0] = 0 \text{ or } 1.$$

- ii) When $b \leq s$, then for all $\beta > 0$

$$\mathbb{P}[W_\infty(\beta) = 0] = 1.$$

- iii) When $s < b$ there is a $\beta_c > 0$ such that $\mathbb{P}[W_\infty(\beta) > 0] = 1$ for $\beta \leq \beta_c$ and $\mathbb{P}[W_\infty(\beta) = 0] = 1$ for $\beta > \beta_c$.

Thus by (ii), when $b \leq s$ the question of weak disorder versus strong disorder is merely that of infinite temperature ($\beta = 0$) versus finite temperature ($\beta > 0$), whereas by (iii), when $s < b$ there is a region of finite temperatures ($\beta \leq \beta_c$) in which weak disorder still holds.

Our goal in this article is to magnify the high temperature regime of the model, $\beta \ll 1$, in the cases $b < s$ and $b = s$ through an examination of the limiting behavior of the random variables $W_n(\beta_n)$ for $n \gg 1$ and some choice of $\beta_n \ll 1$. In other words, the inverse temperature β_n is scaled towards the critical value $\beta_c = 0$ as the size n of the system goes to infinity. The following theorem states our main result in the $b < s$ case.

Theorem 2.1 ($b < s$). *Pick $\hat{\beta} > 0$ and define $\beta_n := \hat{\beta}(b/s)^{\frac{n}{2}}$. As $n \rightarrow \infty$, there is convergence in distribution*

$$W_n(\beta_n) \xrightarrow{\mathcal{L}} L_{\hat{\beta}^2 \frac{s-1}{s-b}}^{b,s}$$

for a family of probability distributions $(L_r^{b,s})_{r>0}$ supported on \mathbb{R}^+ and satisfying the following:

I) $L_r^{b,s}$ has mean 1 and variance $\mathbf{v}_{b,s}(r)$ for a function $\mathbf{v}_{b,s} : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ satisfying

$$\mathbf{v}_{b,s}\left(\frac{s}{b}r\right) = \frac{1}{b}[(1+r)^s - 1] \quad \text{and} \quad \lim_{r \searrow 0} \frac{\mathbf{v}_{b,s}(r)}{r} = 1.$$

II) If $X_r^{(i,j)}$ are independent random variables with distribution $L_r^{b,s}$, then

$$L_{\frac{s}{b}r}^{b,s} \stackrel{d}{=} \frac{1}{b} \sum_{i=1}^b \prod_{1 \leq j \leq s} X_r^{(i,j)}.$$

III) If X_r has distribution $L_r^{b,s}$, then as $r \searrow 0$ there is convergence in law

$$\frac{X_r - 1}{\sqrt{r}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

Remark 2.2. Given $x > 0$ the family of probability distributions $(L_r^{b,s})_{r>0}$ can be used to define a consistent family of random measures $\mu_x^{(n)}$ on paths Γ_n for $n \in \mathbb{N}$. To do this we attach i.i.d. random variables $X_n(a)$ having distribution $L_{x(b/s)^n}^{b,s}$ to edges $a \in E_n$ of the diamond graph D_n . The random probability assigned to a given path $p \in \Gamma_n$ is then

$$\mu_x^{(n)}(p) = \frac{\prod_{a \in p} X_n(a)}{\sum_{p \in \Gamma_n} \prod_{a \in p} X_n(a)},$$

where the products are over the set of edges a lying along the path p . We show that this definition is self-consistent in the proof of Corollary 4.2.

Remark 2.3. In Lemma 4.16 we show that $L_r^{b,s}$ converges weakly to a δ -distribution at 0 as $r \rightarrow \infty$. This is indicative of a continuous transition to strong disorder as $r \rightarrow \infty$. It would be interesting to determine if there are exponential fluctuations for these laws as $r \rightarrow \infty$, that is if by taking a logarithm and then centering and scaling in an appropriate way a non-trivial limiting law appears. This is the situation that was encountered for the limiting law of the partition function in [1], built upon earlier work in [4]. In that case the limiting law was the ubiquitous Tracy-Widom GUE distribution.

Remark 2.4. We emphasize that our methods for showing the existence of the $L_r^{b,s}$ laws and establishing convergence towards them is not based on polynomial chaos expansions, as in [1, 8, 9]. In the polynomial chaos method the partition function is expanded into a sum of polynomials in the disorder variables and collected according to the degree. Convergence follows by Central Limit Theorem type arguments and limit laws for the transition kernels of the underlying random walk, which act as coefficients of the polynomials. The proofs in this paper are entirely different and instead use the recursive structure of the lattice to eliminate the use of the chaos expansion.

Next we state our main result in the $b = s$ case, for which we scale the inverse temperature as $\beta = \hat{\beta}/n$. We feel that this is an interesting regime, because there is a critical point in the behavior of the normalized partition function depending on whether the parameter $\hat{\beta}$ is less than or greater than a cut-off value

$$\kappa_b := \frac{\pi\sqrt{b}}{\sqrt{2}(b-1)}. \quad (2.3)$$

Theorem 2.5 ($b = s$). *As $n \rightarrow \infty$, we have the following behavior for the variance of $W_n(\hat{\beta}/n)$:*

$$\text{Var}\left(W_n\left(\frac{\hat{\beta}}{n}\right)\right) \rightarrow \begin{cases} 0, & 0 \leq \hat{\beta} \leq \kappa_b, \\ \infty, & \hat{\beta} > \kappa_b. \end{cases} \quad (2.4)$$

When $\hat{\beta} < \kappa_b$ the fluctuations of $W_n(\hat{\beta}/n)$ around 1 obey the limit theorem:

$$n^{\frac{1}{2}}\left(W_n\left(\frac{\hat{\beta}}{n}\right) - 1\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, v_b(\hat{\beta})\right), \quad (2.5)$$

where the variance is $v_b(\hat{\beta}) := \hat{\beta}^{\frac{\sqrt{2}}{\sqrt{b}}} \tan\left(\frac{b-1}{\sqrt{2b}}\hat{\beta}\right)$. At the critical value $\hat{\beta} = \kappa_b$ there is a different scaling:

$$\sqrt{\log n}\left(W_n\left(\frac{\kappa_b}{n}\right) - 1\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{6}{b+1}\right). \quad (2.6)$$

Remark 2.6. We will see in Section 5 that the cut-off κ_b is the blow-up point of the solution $\phi_b : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ to the differential equation

$$\frac{d}{dr}\phi_b(r) = \frac{b-1}{2}(\phi_b(r))^2 + \frac{b-1}{b}, \quad \text{with } \phi_b(0) = 0.$$

The solution is explicitly given by $\phi_b(r) = \frac{\sqrt{2}}{\sqrt{b}} \tan\left(\frac{b-1}{\sqrt{2b}}r\right)$. Note that $\kappa_2 = \pi$.

Remark 2.7. For $\hat{\beta} < \kappa_b$ and large n , a simple Taylor expansion shows that the free energy-type quantity $\log(Z_n(\hat{\beta}/n))$ is close to $W_n(\hat{\beta}/n) - 1$ and the convergence (2.5) is equivalent to

$$n^{\frac{1}{2}}\log\left(Z_n\left(\frac{\hat{\beta}}{n}\right)\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, v_b(\hat{\beta})\right).$$

To give a more complete picture of the behavior of $W_n(\beta_n)$ with $\beta_n \searrow 0$ for large n in the $b \neq s$ cases, we state the following theorem.

Theorem 2.8. *Let $\hat{\beta} > 0$ and $\beta_n \searrow 0$.*

i) When $b < s$ we have the following convergences in law for $W_n(\beta_n)$ depending on β_n :

$$W_n(\beta_n) \xrightarrow{\mathcal{L}} \begin{cases} 1, & \beta_n\left(\frac{s}{b}\right)^n \rightarrow 0, \\ L_{\hat{\beta}^2 \frac{s-1}{s-b}}^{b,s}, & \beta_n\left(\frac{s}{b}\right)^n \rightarrow \hat{\beta}, \\ 0, & \beta_n\left(\frac{s}{b}\right)^n \rightarrow \infty. \end{cases}$$

ii) When $b > s$ then $W_n(\beta_n)$ converges to 1. A more refined characterization of $W_n(\beta_n)$ is

$$\frac{1}{\beta_n}(W_n(\beta_n) - 1) \xrightarrow{\mathcal{P}} \sum_{m=1}^{\infty} \frac{1}{b^m} \sum_{a \in V_m} \omega_a, \quad (2.7)$$

where V_m is the set of m^{th} generation vertices.

Remark 2.9. In (i) of the above theorem, the convergence of $W_n(\beta_n)$ to zero in the $b < s$ case when $\beta_n(b/s)^{n/2} \rightarrow \infty$ follows from the homogeneous environment tilting method in the proofs from Sections 5.2-5.6 of [27].

Remark 2.10. In the $s < b$ case, since there is a region of finite temperature in which weak disorder holds, it would be interesting to find an intermediate disorder regime by scaling towards the critical value of β_c as the system size grows, that is, to consider

$$W_n(\beta_c + \beta_n)$$

for an appropriate sequence $\beta_n \searrow 0$. Perturbations of this type are difficult to study, however, since they typically require fine properties of the Gibbs measure $\mu_{\beta_c, n}^{(\omega)}$ at criticality. For directed polymers on trees results of this type were obtained in [3], based on an earlier analysis of the Gibbs measure at criticality in [24].

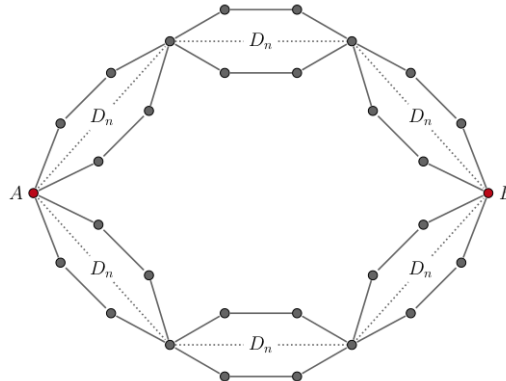
3 Basic constructions, notation, and heuristics

3.1 More on diamond graphs

Now we will give a more in-depth description of the diamond graph structure. The set of copies of D_n found in D_{m+n} has a canonical one-to-one correspondence with the edges of D_m . We will use the following notation:

- The symbol D_n for the n^{th} diamond graph also denotes the set of vertices on the graph, excluding the roots A and B .
- For $k \leq n$, let $V_k \subset D_n$ denote the set of k^{th} generation vertices in D_n , in other words, those vertices that appear in D_k but not in D_{k-1} .
- E_n denotes the set of edges on the graph D_n .
- For $k \leq n$, $G_{k,n}$ denotes the set of copies of D_k found in D_n . An element $g \in G_{k,n}$ is also regarded as a subset of the set of vertices D_n .

By construction of D_n the elements in $G_{k,n}$ are in one-to-one correspondence with elements in E_{n-k} .



Copies of D_n embedded in D_{n+m} .

The edges of D_1 can be labeled by $(i, j) \in \{1, \dots, b\} \times \{1, \dots, s\}$, where (i, j) refers to the j^{th} segment of the i^{th} branch. More generally, the set of edges E_n in D_n can be identified with

$$E_n \equiv (\{1, \dots, b\} \times \{1, \dots, s\})^n.$$

The correspondence follows from the recursive procedure for the construction of D_n :

(Encodes an element of E_1 in the local copy of D_1)

$$\underbrace{(i_1, j_1) \times \dots \times (i_{n-1}, j_{n-1})}_{\text{set of copies of } D_1 \text{ in } D_n} \times \overbrace{(i_n, j_n)}$$

(Encodes an element of $G_{1,n} \equiv E_{n-1}$, the set of copies of D_1 in D_n)

Similarly, elements in $G_{k,n}$ can be labeled by $(\{1, \dots, b\} \times \{1, \dots, s\})^{n-k}$. The elements of $G_{k-1,n}$ that are contained in $g \in G_{k,n}$ will be labeled by

$$g \times (i, j),$$

where $i \in \{1, \dots, b\}$, and $j \in \{1, \dots, s\}$.

The k^{th} generation vertices V_k are identified with the set

$$G_{n-k+1,n} \times (\{1, \dots, b\} \times \{1, \dots, s-1\}).$$

For $g \in G_{k,n}$, we will label the generation $n-k+1$ vertices that are contained in g by

$$g \oslash (i, j),$$

where $i \in \{1, \dots, b\}$ and $j \in \{1, \dots, s-1\}$.

We have the following basic combinatorial observations regarding these graphs:

- i) The graph D_n has $b^n s^n$ edges, i.e., $|E_n| = b^n s^n$. Since the subgraphs $G_{k,n}$ and the edges E_{n-k} are in one-to-one correspondence, this implies that $|G_{k,n}| = b^{n-k} s^{n-k}$.
- ii) The graph D_n has $b^k s^{k-1}(s-1)$ vertices of generation k when $k \leq n$, i.e., $|V_k| = b^k s^{k-1}(s-1)$.
- iii) An element $g \in G_{k,n}$ contains $b^{j-n+k} s^{j-n+k-1}(s-1)$ vertices of generation $j \in (n-k, n]$.
- iv) There are $b^{\frac{s^n-1}{s-1}}$ paths between the root vertices A and B in D_n , i.e., $|\Gamma_n| = b^{\frac{s^n-1}{s-1}}$.

3.2 Definitions and recursive relations for statistical quantities

Recall that $G_{k,n}$ refers to the collection of copies of D_k found as subgraphs of D_n . For $g \in G_{k,n}$, we define $W_n(\beta; g)$ in analogy to $W_n(\beta)$ except restricted to the subgraph g . To be clear, $W_n(\beta; g)$ does not depend on the random variables at the roots of the subgraph g . The following lemma is essentially a restatement of (2.2) in the subgraph notation:

Lemma 3.1. *For $g \in G_{k,n}$, the random variables $W_n(\beta; g)$ satisfy the recursive relation*

$$W_n(\beta; g) = \frac{1}{b} \sum_{i=1}^b \prod_{1 \leq j \leq s} W_n(\beta; g \times (i, j)) \prod_{1 \leq j \leq s-1} E_{g \oslash (i, j)}(\beta) \quad (3.1)$$

with $W_n(\beta; g) := 1$ for $g \in G_{0,n}$.

From this lemma it is straightforward to derive a recursive relation for the variance of the W_n variables.

Corollary 3.2. *For $g \in G_{k,n}$, the variance $\sigma_k(\beta) := \text{Var}[W_n(\beta; g)]$ satisfies the recursive relation in $k \in \mathbb{N}$*

$$\sigma_{k+1}(\beta) = \frac{1}{b} \left[(1 + \sigma_k(\beta))^s e^{(s-1)[\lambda(2\beta) - 2\lambda(\beta)]} - 1 \right].$$

3.3 Heuristic motivation

We now offer a short heuristic analysis motivating the scalings $\beta = \widehat{\beta}(b/s)^{n/2}$ and $\beta = \widehat{\beta}/n$ for the respective cases $b < s$ and $b = s$. We begin by naively linearizing the expression (2.1) for the normalized partition function to obtain

$$W_n(\beta_n) \approx \frac{1}{|\Gamma_n|} \sum_{p \in \Gamma_n} \prod_{a \in p} (1 + \beta_n \omega_a),$$

which is intuitively justified for small β_n . The right side above can be expanded into the form

$$= 1 + \sum_{m \geq 1} \sum_{\substack{a_1, \dots, a_m \in D_n \\ a_i \neq a_j}} \beta_n^m P(a_1, \dots, a_m) \omega_{a_1} \cdots \omega_{a_m}, \quad (3.2)$$

where $P(a_1, \dots, a_m)$ is the probability that the vertices $a_1, \dots, a_m \in D_n$ lie on a random path $p \in \Gamma_n$. We seek a choice of β_n such that $W_n(\beta_n)$ converges in law to a nontrivial limit with large n . A reasonable method for finding a candidate for β_n is to determine the appropriate scaling for the first-order term (i.e. the $m = 1$ term) of (3.2), which is merely a sum of i.i.d. random variables, and thus characterized by the central limit theorem. This method of choosing β_n turns out to work for $b < s$ but to be misleading in the case of $b = s$. The variance of the first-order term in (3.2) has the form

$$\begin{aligned} \text{Var} \left(\beta_n \sum_{a \in D_n} P(a) \omega_a \right) &= \beta_n^2 \sum_{a \in D_n} (P(a))^2 = \beta_n^2 \sum_{k=1}^n (\# \text{ of } k^{\text{th}} \text{ generation vertices}) b^{-2k} \\ &= \beta_n^2 \sum_{k=1}^n \frac{s-1}{s} \left(\frac{s}{b} \right)^k \\ &\approx \beta_n^2 \frac{s-1}{s} \cdot \begin{cases} \left(\frac{s}{b} \right)^n \left(1 - \frac{b}{s} \right)^{-1}, & b < s, \\ n, & b = s. \end{cases} \end{aligned} \quad (3.3)$$

Thus, by the central limit theorem, the first-order term from (3.2) converges in law to a Gaussian when $\beta_n = (b/s)^{n/2}$ and $\beta_n = n^{-1/2}$ in the $b < s$ and $b = s$ cases, respectively.

Now we perform the same exercise with the second-order term in (3.2) to check whether the above scalings are still appropriate. For $a_1, a_2 \in D_n$, the probability $P(a_1, a_2)$ has the form

$$P(a_1, a_2) = \epsilon(a_1, a_2) \exp \{ \mathfrak{l}(a_1, a_2) - \mathfrak{g}(a_1) - \mathfrak{g}(a_2) \},$$

where $\epsilon(a_1, a_2)$ is 0 when no directed path passes through both of the vertices a_1 and a_2 , $\mathfrak{g}(a)$ is the generation of the vertex, and $\mathfrak{l}(a_1, a_2)$ is the smallest k such that $a_1, a_2 \in g$ for some $g \in G_{n-k, n}$. The variance of the second-order term has the form

$$\text{Var} \left(\beta_n^2 \sum_{\substack{a_1, a_2 \in D_n \\ a_1 \neq a_2}} P(a_1, a_2) \omega_{a_1} \omega_{a_2} \right) = \beta_n^4 \sum_{\substack{a_1, a_2 \in D_n \\ a_1 \neq a_2}} (P(a_1, a_2))^2 = \beta_n^4 \sum_{k=1}^n \sum_{\substack{a_1, a_2 \in D_n \\ k = \mathfrak{l}(a_1, a_2)}} (P(a_1, a_2))^2.$$

The self-similar structure of the diamond graphs allows the above to be written as

$$\beta_n^4 \sum_{k=1}^n b^{-2j} |G_{n-k, n}| \left(\frac{b(s-1)(s-2)}{2} + bs(s-1) \sum_{a \in D_{n-k}} (P(a))^2 + \frac{bs(s+1)}{2} \left(\sum_{a \in D_{n-k}} (P(a))^2 \right)^2 \right).$$

The three terms above correspond to the respective cases $\mathfrak{g}(a_1) = \mathfrak{g}(a_2) = k$, $\mathfrak{g}(a_1) < \mathfrak{g}(a_2) = k$ (or vice versa), and $\mathfrak{g}(a_1), \mathfrak{g}(a_2) < k$. The third term is dominant, and we can apply that $|G_{n-k,n}| = (bs)^k$ and our analysis in (3.3) to get that the above is approximately

$$\approx \beta_n^4 \sum_{k=1}^n \left(\frac{s}{b}\right)^k \frac{bs(s+1)}{2} \left(\frac{s-1}{s}\right)^2 \cdot \begin{cases} \left(\frac{s}{b}\right)^{2(n-k)} \left(1 - \frac{b}{s}\right)^{-2}, & b < s, \\ (n-k)^2, & b = s. \end{cases}$$

Finally, by standard summation formulas, this is equal to

$$= \beta_n^4 \frac{(s+1)(s-1)^2}{2} \cdot \begin{cases} \frac{\left(\frac{b}{s}\right)^2}{\left(1 - \frac{b}{s}\right)^3} \left(\frac{s}{b}\right)^{2n} + O\left(\left(\frac{s}{b}\right)^n\right), & b < s, \\ \frac{1}{3}n^3 + O(n^2), & b = s. \end{cases}$$

For $b < s$ the choice $\beta_n = (b/s)^{n/2}$ again yields a variance that converges for the second-order term, as would be expected. In contrast, for $b = s$ the previous choice of $\beta_n = n^{-1/2}$ causes the variance of the second-order term to blow up with large n . Notice, however, that if β_n is scaled in proportion to n^{-1} in the $b = s$ case, then the first- and second-order terms both have variances of the same order, $\propto n^{-1}$. This suggests that $\sqrt{n}(W_n(\hat{\beta}/n) - 1)$ is an interesting rescaled version of the normalized partition function to study when $b = s$. It is possible that there is a nontrivial limit for $W_n(\hat{\beta}/\sqrt{n})$ in the $b = s$ case, but the limiting distribution will not have finite variance and we do not explore this possibility here.

4 A limit theorem in the $b < s$ case

In this section we prove Theorem 2.1. We first work on a slightly simpler problem in which random variables are placed on the edges of the diamond graph rather than the vertices. We use a coupling of Gaussian random variables across the different generations of the lattice to prove that a certain linearized version of the partition function is a Cauchy sequence, thereby establishing the existence of the $L_r^{b,s}$ laws. We then build on this result to show that the limit laws are universal, and finally translate these results from the edge case to the vertex case.

4.1 The edge problem

The proof of the following lemma is below.

Lemma 4.1. *There exists a family of probability distributions $(L_r^{b,s})_{r>0}$ satisfying the properties (I)-(III) listed in Theorem 2.1.*

The family of measures $(L_r^{b,s})_{r>0}$ from Lemma 4.1 can be used to construct a family of random probability measures on the set of directed paths Γ_n whose laws are consistent for all $n \in \mathbb{N}$. For $k < n$ the path set Γ_k on D_k is canonically identified with a partition of Γ_n by relating paths that pass through the same collection of k^{th} generation vertices. This consistency property means that $(L_r^{b,s})_{r>0}$ can be used to construct random measures on the set of directed paths, Γ_∞ , on the continuum diamond lattice D_∞ , which we will not discuss further in this article. Given $a \in E_n$ and $p \in \Gamma_n$, we will write $a \triangleleft p$ to mean that the path p passes through the edge a .

Corollary 4.2. *Pick $x > 0$. Let $\mu_k^{b,s}$ denote a random probability measure on Γ_k such that for an independent family of random variables $\{X(a)\}_{a \in E_k}$ with distribution $L_{x(b/s)^k}^{b,s}$ a path $p \in \Gamma_k$ is assigned probability*

$$\mu_k^{b,s}(p) = \frac{\prod_{a \triangleleft p} X(a)}{\sum_{q \in \Gamma_k} \prod_{a \triangleleft q} X(a)},$$

where the products are over the set of edges, a , lying along the path $p \in \Gamma_k$. For $n > k$, let S_p be the set of paths on D_n that pass through the same set of k^{th} generation vertices as p . Then the measures $\mu_k^{b,s}$ are consistent for all $k \in \mathbb{N}$ in the sense that the family of random variables $\{\mu_n^{b,s}(S_p)\}_{p \in \Gamma_k}$ has the same law as $\{\mu_k^{b,s}(p)\}_{p \in \Gamma_k}$.

Proof. For $g \in G_{0,n} \equiv E_n$, let $X(g)$ be independent random variables with distribution $L_{x(b/s)^n}^{b,s}$. We can define independent random variables $X(g)$ labeled by $g \in G_{n-k,n} \equiv E_k$ through inductive use of the recursive relation

$$X(g) = \frac{1}{b} \sum_{i=1}^b \prod_{1 \leq j \leq s} X(g \times (i, j)).$$

The $X(g)$, $g \in E_k$ have distribution $L_{x(b/s)^k}^{b,s}$ as a consequence of property (II) listed in Theorem 2.1. Notice that

$$\mu_n^{b,s}(S_p) \stackrel{d}{=} \frac{\sum_{q \in S_p} \prod_{a \triangleleft q} X(a)}{\sum_{q \in \Gamma_n} \prod_{a \triangleleft q} X(a)} = \frac{\prod_{g \triangleleft p} X(g)}{\sum_{q \in \Gamma_k} \prod_{g \triangleleft q} X(g)} \stackrel{d}{=} \mu_k^{b,s}(p).$$

Thus the laws of the measures are consistent. This equality in law generalizes to the families $\{\mu_n^{b,s}(S_p)\}_{p \in \Gamma_k}$, $\{\mu_k^{b,s}(p)\}_{p \in \Gamma_k}$. □

Definition 4.3. Given numbers $X_g \in \mathbb{R}$ labeled by $g \in E_n \equiv G_{0,n}$, we define

- $\mathbf{W}_n(\{X_g\}_{g \in E_n})$ to be equal to $\widehat{W}_n(D_n)$ for the array of real numbers $\widehat{W}_n(g)$, $g \in G_{k,n}$ determined by the recursive relation

$$\widehat{W}_n(g) = \frac{1}{b} \sum_{i=1}^b \prod_{1 \leq j \leq s} \widehat{W}_n(g \times (i, j)) \quad (4.1)$$

with initial condition $\widehat{W}(g) = X_g$ for $g \in G_{0,n}$.

- $\overline{\mathbf{W}}_n(\{X_g\}_{g \in E_n})$ analogously to the above through the linearized recursive relation

$$\widehat{W}_n(g) = 1 + \frac{1}{b} \sum_{i=1}^b \sum_{1 \leq j \leq s} (\widehat{W}_n(g \times (i, j)) - 1). \quad (4.2)$$

Remark 4.4. If $\{X_a\}_{a \in E_n}$ is an array of real numbers and $k < n$, then

$$\mathbf{W}_n(\{X_a\}_{a \in E_n}) = \mathbf{W}_k(\{\mathbf{W}_{n-k}(\{X_a\}_{a \triangleleft g})\}_{g \in E_k}),$$

where we have abused notation by using the one-to-one correspondence between E_k and $G_{n-k,n}$ and that $G_{n-k,n}$ defines a partition of E_n to classify each edge $a \in E_n$ as a member of an edge g in E_k .

Remark 4.5. Once we have proved Theorem 2.1 we will have the following: if $\{\mathbf{x}_g\}_{g \in E_k}$ is an array of independent random variables having distribution $L_x^{b,s}$, then $\mathbf{W}_k(\{\mathbf{x}_g\}_{g \in E_k})$ has law $L_{x(s/b)^k}^{b,s}$.

We will work towards a proof of the following limit theorem.

Theorem 4.6 (Edge limit theorem). *For $n \in \mathbb{N}$, let $X_g^{(n)}$ be an array of positive i.i.d. random variables labeled by $g \in E_n$ with mean 1 and variance $x_n \in \mathbb{R}^+$. In addition, we assume that*

$$\lim_{n \rightarrow \infty} (s/b)^n x_n = x \quad \text{and} \quad \lim_{\lambda \nearrow \infty} \sup_{n \in \mathbb{N}} (sb)^n \mathbb{P} \left[|X_g^{(n)} - 1|^2 > \lambda (b/s)^n \right] = 0$$

for some $x > 0$. Then there is weak convergence as $n \rightarrow \infty$ given by

$$\mathbf{W}_n(\{X_g^{(n)}\}_{g \in E_n}) \xrightarrow{\mathcal{L}} L_x^{b,s},$$

where the family of probability distributions $(L_r^{b,s})_{r \geq 0}$ satisfies the properties listed in Theorem 2.1.

Definition 4.7. Define $\widehat{M} : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ by

$$\widehat{M}(x) := \frac{1}{b} [(1+x)^s - 1].$$

Remark 4.8. Let X_g be an array of positive i.i.d. random variables labeled by $g \in E_k$ having mean 1 and variance $x \in \mathbb{R}^+$. The k -fold composition of the map \widehat{M} yields

$$\text{Var}(\mathbf{W}_k(\{X_g\}_{g \in E_k})) = \widehat{M}^k(x).$$

We will also make repeated use of the simple formula

$$\text{Var}(\overline{\mathbf{W}}_k(\{X_g\}_{g \in E_k})) = \left(\frac{s}{b}\right)^k x.$$

Next we collect some estimates on \widehat{M} . The proof of Lemma 4.9 is placed in Section 8.1.

Lemma 4.9. *For any $\lambda > 0$, there is a $C > 0$ such that the following inequalities hold for all $0 \leq x \leq \lambda$, $N \in \mathbb{N}$, and $N \geq n$:*

- i) $\widehat{M}^{N-n}(x(b/s)^N) \leq Cx(b/s)^n$,
- ii) $\widehat{M}^{N-n}(x(b/s)^N) - x(b/s)^n \leq Cx^2(b/s)^{2n}$, and
- iii) $\frac{d}{dx} [\widehat{M}^n(x(b/s)^n)] \leq C$.
- iv) For unrestricted $x \geq 0$, we still have a bound of the form

$$\frac{d}{dx} [\widehat{M}^n(x(b/s)^n)] \leq \frac{s}{b} \exp \left\{ \frac{s-1}{1-\frac{b}{s}} \widehat{M}^n(x(b/s)^n) \right\}.$$

Corollary 4.10 (Limiting variance function). *There is an increasing function $\mathbf{v}_{b,s} : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ such that for all $x \in \mathbb{R}^+$*

I) As $n \rightarrow \infty$,

$$\widehat{M}^n(x(b/s)^n) \nearrow \mathbf{v}_{b,s}(x).$$

II) $\mathbf{v}_{b,s}$ satisfies the relation

$$\mathbf{v}_{b,s}\left(\frac{s}{b}x\right) = \widehat{M}(\mathbf{v}_{s,b}(x)).$$

III) The derivative of $\mathbf{v}_{b,s}$ satisfies $\lim_{x \searrow 0} \mathbf{v}'_{b,s}(x) = 1$ and

$$\mathbf{v}'_{b,s}(x) = \prod_{j=1}^{\infty} \left(1 + \mathbf{v}_{b,s}(x(b/s)^j)\right)^{s-1} = \prod_{j=1}^{\infty} \left(1 + \widehat{M}^{-j}(\mathbf{v}_{b,s}(x))\right)^{s-1}.$$

Proof. Properties (II) and (III) follow from the existence of the limit (I). To see that the sequence $\widehat{M}^n(x(b/s)^n)$ is Cauchy, notice that for $N > n \gg 1$

$$\begin{aligned} \widehat{M}^N(x(b/s)^N) - \widehat{M}^n(x(b/s)^n) &= \widehat{M}^n\left(\widehat{M}^{N-n}(x(b/s)^N)\right) - \widehat{M}^n(x(b/s)^n) \\ &\leq \left[\widehat{M}^{N-n}(x(b/s)^N) - x(b/s)^n\right] \frac{d}{dy} \widehat{M}^n(y) \Big|_{y=\widehat{M}^{N-n}(x(b/s)^N)} \\ &\leq Cx^2(b/s)^{2n} \frac{d}{dy} \widehat{M}^n(y) \Big|_{y=c(b/s)^n} \\ &\leq C^2x^2(b/s)^{2n}. \end{aligned}$$

In the first inequality above, we have used that the derivative of \widehat{M} is an increasing function. The second inequality uses (ii) and (i) of Lemma 4.9 for the first and second factors, respectively. The third inequality uses (iii) of the same lemma. □

Lemma 4.11. For each $n \in \mathbb{N}$, let $Y_g^{(n)}$ and $Z_g^{(n)}$ be arrays of i.i.d. random variables labeled by $g \in E_n$ such that $Y_g^{(n)}$ and $Z_g^{(n)}$ are uncorrelated, $\mathbb{E}[Y_g^{(n)}] = 1$, $\mathbb{E}[Z_g^{(n)}] = 0$, and for large $n > 0$

$$\bar{y} := \sup_{n \in \mathbb{N}} (s/b)^n \text{Var}(Y_g^{(n)}) < \infty \quad \text{and} \quad \text{Var}(Z_g^{(n)}) = o((b/s)^n).$$

Then, the following holds:

i) There is convergence to zero in probability as $n \rightarrow \infty$ given by

$$\left| \mathbf{W}_n(\{Y_g^{(n)} + Z_g^{(n)}\}_{g \in E_n}) - \mathbf{W}_n(\{Y_g^{(n)}\}_{g \in E_n}) \right| \xrightarrow{\mathcal{P}} 0.$$

ii) For $g \in E_n$ and $N > n$, the random variables $\mathbf{W}_{N-n}(\{Y_a^{(N)}\}_{a \triangleleft g}) - \overline{\mathbf{W}}_{N-n}(\{Y_a^{(N)}\}_{a \triangleleft g})$ and $\overline{\mathbf{W}}_{N-n}(\{Y_a^{(N)}\}_{a \triangleleft g})$ are uncorrelated, and there is a $C > 0$ such that for all $N, n \in \mathbb{N}$ and $Y_a^{(N)}$

$$\text{Var}\left(\mathbf{W}_{N-n}(\{Y_g^{(N)}\}_{a \triangleleft g}) - \overline{\mathbf{W}}_{N-n}(\{Y_g^{(N)}\}_{a \triangleleft g})\right) \leq C\bar{y}^2(b/s)^{2n}.$$

Proof. Part (i): Define $y_n := (s/b)^n \text{Var}(Y_g^{(n)})$ and $x_n := (s/b)^n \text{Var}(Y_g^{(n)} + Z_g^{(n)})$. The random variables $\mathbf{W}_n(\{Y_g^{(n)} + Z_g^{(n)}\}_{g \in E_n}) - \mathbf{W}_n(\{Y_g^{(n)}\}_{g \in E_n})$ and $\mathbf{W}_n(\{Y_g^{(n)}\}_{g \in E_n})$ are uncorrelated and thus we have first equality below.

$$\begin{aligned} \mathbb{E} \left[\left| \mathbf{W}_n(\{Y_g^{(n)} + Z_g^{(n)}\}_{g \in E_n}) - \mathbf{W}_n(\{Y_g^{(n)}\}_{g \in E_n}) \right|^2 \right] &= \text{Var}\left(\mathbf{W}_n(\{Y_g^{(n)} + Z_g^{(n)}\}_{g \in E_n})\right) - \text{Var}\left(\mathbf{W}_n(\{Y_g^{(n)}\}_{g \in E_n})\right) \\ &= \widehat{M}^n(x_n(b/s)^n) - \widehat{M}^n(y_n(b/s)^n) \\ &\leq \frac{d}{dx} \left[\widehat{M}^n(x(b/s)^n) \right] \Big|_{x=x_n} (x_n - y_n). \end{aligned}$$

The second equality follows from Remark 4.8 and the inequality uses that \widehat{M} has increasing derivative. By our assumptions, x_n is bounded for all n and $|x_n - y_n| \rightarrow 0$. Hence, we can apply (i) and (iii) of Lemma 4.9 to show that the above converges to zero with large n .

Part (ii): That the stated variables are uncorrelated follows from the definition of \mathbf{W} and $\overline{\mathbf{W}}$. Since they are uncorrelated we have

$$\begin{aligned} \text{Var}\left(\mathbf{W}_{N-n}(\{Y_a^{(N)}\}_{a \triangleleft g}) - \overline{\mathbf{W}}_{N-n}(\{Y_a^{(N)}\}_{a \triangleleft g})\right) \\ = \text{Var}\left(\mathbf{W}_{N-n}(\{Y_a^{(N)}\}_{a \triangleleft g})\right) - \text{Var}\left(\overline{\mathbf{W}}_{N-n}(\{Y_a^{(N)}\}_{a \triangleleft g})\right) \\ = \widehat{M}^{N-n}(y_N(b/s)^N) - y_N(b/s)^n \\ \leq C y_N^2(b/s)^{2n} \leq C \bar{y}^2(b/s)^{2n}, \end{aligned}$$

where the first inequality holds by part (ii) of Lemma 4.9. □

The proof of Lemma 4.1 is a warm-up for the proof of Theorem 2.1, and we will refer back to some of the constructions in the proof.

Proof of Lemma 4.1. Fix $r > 0$. For $n \in \mathbb{N}$ let $\{\widehat{\omega}_g\}_{g \in E_n}$ be arrays of independent random variables with distribution $\mathcal{N}(0, r(b/s)^n)$ and satisfying a hierarchical consistency relation:

$$\widehat{\omega}_g = \frac{1}{b} \sum_{i=1}^b \sum_{1 \leq j \leq s} \widehat{\omega}_{g \times (i,j)}.$$

The above can be easily constructed using, for instance, two independent standard Brownian motions $\mathbf{B}^{(1)}$ and $\mathbf{B}^{(2)}$ and the fact that E_n has a canonical one-to-one correspondence with $(\{1, \dots, b\} \times \{1, \dots, s\})^n$ to identify the $\widehat{\omega}_g$, $g \in E_n$ with variables of the form

$$b^n \left(\mathbf{B}_{\frac{i+1}{b^n}}^{(1)} - \mathbf{B}_{\frac{i}{b^n}}^{(1)} \right) \left(\mathbf{B}_{\frac{j+1}{s^n}}^{(2)} - \mathbf{B}_{\frac{j}{s^n}}^{(2)} \right),$$

where $0 \leq i \leq b^n - 1$ and $0 \leq j \leq s^n - 1$. This construction has the useful property that for $N > n$ and $g \in G_{n,N}$,

$$1 + \widehat{\omega}_g = \overline{\mathbf{W}}_{N-n}(\{1 + \widehat{\omega}_a\}_{a \triangleleft g}). \quad (4.3)$$

We will show that there is convergence in probability

$$\mathbf{W}_n(\{1 + \widehat{\omega}_g\}_{g \in E_n}) \xrightarrow{\mathcal{P}} (\text{Some limit}). \quad (4.4)$$

It is enough to show that $\mathbf{W}_n(\{\widehat{\omega}_g\}_{g \in E_n})$ is Cauchy in probability. Notice that for $N > n$ the equality (4.3) implies the first equality below

$$\begin{aligned} \mathbf{W}_N(\{1 + \widehat{\omega}_a\}_{a \in E_N}) - \mathbf{W}_n(\{1 + \widehat{\omega}_g\}_{g \in E_n}) \\ = \mathbf{W}_N(\{1 + \widehat{\omega}_a\}_{a \in E_N}) - \mathbf{W}_n(\{\overline{\mathbf{W}}_{N-n}(\{1 + \omega_a\}_{a \triangleleft g})\}_{g \in E_n}) \\ = \mathbf{W}_n(\{Y_g^{(n)} + Z_g^{(n)}\}_{g \in E_n}) - \mathbf{W}_n(\{Y_g^{(n)}\}_{g \in E_n}), \end{aligned}$$

where for $g \in E_n \equiv G_{N-n,N}$

$$Y_g^{(N)} := \overline{\mathbf{W}}_{N-n}(\{1 + \widehat{\omega}_a\}_{a \triangleleft g}) \quad \text{and} \quad Z_g^{(N)} := \mathbf{W}_{N-n}(\{1 + \widehat{\omega}_a\}_{a \triangleleft g}) - \overline{\mathbf{W}}_{N-n}(\{1 + \widehat{\omega}_a\}_{a \triangleleft g}).$$

In the above, we have used the one-to-one correspondence between E_n and $G_{N-n,N}$ to classify edges $a \in E_N$ as elements g in E_n (since $G_{N-n,N}$ defines a partition of E_N). Notice that $Y_g^{(n)}$ and $Z_g^{(n)}$ are uncorrelated and

$$\text{Var}(Y_g^{(n)}) = r(b/s)^n \quad \text{and} \quad \text{Var}(Z_g^{(n)}) = o((b/s)^n),$$

where the order equality follows from part (ii) of Lemma 4.11. Hence we can apply part (i) of Lemma 4.11 to conclude that the difference between $\mathbf{W}_n(\{Y_g^{(n)} + Z_g^{(n)}\}_{g \in E_n})$ and $\mathbf{W}_n(\{Y_g^{(n)}\}_{g \in E_n})$ converges in probability to zero with large n .

The properties (I)-(III) of Theorem 2.1 follow closely from the above limit construction. For (II), notice that as $n \rightarrow \infty$ for each $(i, j) \in \{1, \dots, b\} \times \{1, \dots, s\}$ our result (4.4) implies

$$\mathbf{W}_{n-1}(\{1 + \hat{\omega}_{g \times (i,j)}\}_{g \in E_n}) \xrightarrow{\mathcal{L}} X_{(i,j)},$$

where the $X_{(i,j)}$ are independent random variables with law $L_{rb/s}^{b,s}$. Moreover,

$$\frac{1}{b} \sum_{i=1}^b \prod_{1 \leq j \leq s} \mathbf{W}_{n-1}(\{1 + \hat{\omega}_{g \times (i,j)}\}_{g \in E_n}) = \mathbf{W}_n(\{1 + \hat{\omega}_g\}_{g \in E_n}) \xrightarrow{\mathcal{L}} L_r^{b,s}, \quad (4.5)$$

where the convergence in law holds by another application of the limit result above. Thus (II) holds. For (III), we write

$$\begin{aligned} & \frac{\mathbf{W}_n(\{1 + \hat{\omega}_g\}_{g \in E_n}) - 1}{\sqrt{r}} \\ &= \underbrace{\frac{\mathbf{W}_n(\{1 + \hat{\omega}_g\}_{g \in E_n}) - \overline{\mathbf{W}}_n(\{1 + \hat{\omega}_g\}_{g \in E_n})}{\sqrt{r}}}_{\xrightarrow{\mathcal{P}} 0 \text{ as } r \searrow 0} + \underbrace{\frac{\overline{\mathbf{W}}_n(\{1 + \hat{\omega}_g\}_{g \in E_n}) - 1}{\sqrt{r}}}_{\stackrel{d}{=} \mathcal{N}(0,1)}, \end{aligned}$$

where the convergence as $r \searrow 0$ of the first term is uniform in large n by part (ii) of Lemma 4.11. The normal distribution of the second term follows from (4.2) being a linear recursion and the initial variables being normal. \square

Proof of Theorem 4.6. Fix some $0 < \epsilon \ll 1$ and $n \in \mathbb{N}$ such that $x(b/s)^n < \epsilon$. For $N > n$, let $\{\hat{\omega}_a\}_{a \in E_N}$ be an array of normal random variables defined as in the proof of Lemma 4.1 for r replaced by x . By the method in the proof of Lemma 4.1, we can construct an array $\{\mathbf{x}_g\}_{g \in E_n}$ of independent random variables with distribution $L_{x(b/s)^n}^{b,s}$ through limits in probability as $N \rightarrow \infty$ of the form

$$\mathbf{W}_{N-n}(\{1 + \hat{\omega}_a\}_{a \triangleleft g}) \xrightarrow{\mathcal{P}} \mathbf{x}_g, \quad (4.6)$$

where $\{1 + \hat{\omega}_a\}_{a \triangleleft g}$ refers to the subarray of $\{1 + \hat{\omega}_a\}_{a \in E_N}$ comprised of all edges $a \in E_N$ lying on the subgraph $g \in G_{N-n,N} \equiv E_n$. Throughout this proof we will freely identify the sets $G_{N-n,N}$ and E_n , allowing us to treat elements g of E_n as containing sets of edges from E_N . Notice that $\mathbf{W}_n(\{\mathbf{x}_g\}_{g \in E_n})$ has law $L_x^{b,s}$ by Remark 4.5. We will show that for arbitrary $F : \mathbb{R} \rightarrow \mathbb{R}$ with bounded derivative that in the limit $N \rightarrow \infty$

$$\mathbb{E}\left[F\left(\mathbf{W}_N(\{X_a^{(N)}\}_{a \in E_N})\right)\right] \longrightarrow \mathbb{E}\left[F\left(\mathbf{W}_n(\{\mathbf{x}_g\}_{g \in E_n})\right)\right]. \quad (4.7)$$

The above guarantees the weak convergence of $\mathbf{W}_n(\{X_a^{(n)}\}_{a \in E_n})$ to $L_x^{b,s}$.

The difference between the terms in (4.7) is bounded through the triangle inequality as follows:

$$\begin{aligned}
& \left| \mathbb{E} \left[F \left(\mathbf{W}_N(\{X_a^{(N)}\}_{a \in E_N}) \right) \right] - \mathbb{E} \left[F \left(\mathbf{W}_n(\{\mathbf{x}_g\}_{g \in E_n}) \right) \right] \right| \\
& \leq \left| \mathbb{E} \left[F \left(\mathbf{W}_N(\{X_a^{(N)}\}_{a \in E_N}) \right) \right] - \mathbb{E} \left[F \left(\mathbf{W}_n(\{\overline{\mathbf{W}}_{N-n}(\{X_a^{(N)}\}_{a \triangleleft g})\}_{g \in E_n}) \right) \right] \right| \\
& \quad + \left| \mathbb{E} \left[F \left(\mathbf{W}_n(\{\overline{\mathbf{W}}_{N-n}(\{X_a^{(N)}\}_{a \triangleleft g})\}_{g \in E_n}) \right) \right] - \mathbb{E} \left[F \left(\mathbf{W}_n(\{1 + \widehat{\omega}_g\}_{g \in E_n}) \right) \right] \right| \\
& \quad + \left| \mathbb{E} \left[F \left(\mathbf{W}_n(\{1 + \widehat{\omega}_g\}_{g \in E_n}) \right) \right] - \mathbb{E} \left[F \left(\mathbf{W}_n(\{\mathbf{W}_{N-n}(\{1 + \widehat{\omega}_a\}_{a \triangleleft g})\}_{g \in E_n}) \right) \right] \right| \\
& \quad + \left| \mathbb{E} \left[F \left(\mathbf{W}_n(\{\mathbf{W}_{N-n}(\{1 + \widehat{\omega}_a\}_{a \triangleleft g})\}_{g \in E_n}) \right) \right] - \mathbb{E} \left[F \left(\mathbf{W}_n(\{\mathbf{x}_g\}_{g \in E_n}) \right) \right] \right|. \quad (4.8)
\end{aligned}$$

It is sufficient to show that the terms above are each bounded by a constant multiple of $\epsilon > 0$ in the limit $N \rightarrow \infty$. The last term on the right side of (4.8) converges to zero as $N \rightarrow \infty$ by (4.6) and since $\mathbf{W}_n(\{a_g\}_{g \in E_n})$ is a continuous function (a multivariable polynomial) of the array $\{a_g\}_{g \in E_n}$. We will treat the first three terms on the right side of (4.8) by proving the following statements:

(I) There is a $C > 0$ such that for all $N, n \in \mathbb{N}$ with $N > n$

$$\text{Var} \left(\mathbf{W}_N(\{X_g^{(N)}\}_{g \in E_N}) - \mathbf{W}_n(\{\overline{\mathbf{W}}_{N-n}(\{X_a^{(N)}\}_{a \triangleleft g})\}_{g \in E_n}) \right) \leq Cx^2(b/s)^n.$$

Consequently, the first term on the right side of (4.8) is bounded by $\epsilon\sqrt{C} \sup_{x \in \mathbb{R}} |F'(x)|$.

(II) For any fixed n , as $N \rightarrow \infty$

$$\mathbb{E} \left[F \left(\mathbf{W}_n(\{\overline{\mathbf{W}}_{N-n}(\{X_a^{(N)}\}_{a \triangleleft g})\}_{g \in E_n}) \right) \right] \longrightarrow \mathbb{E} \left[F \left(\mathbf{W}_n(\{1 + \widehat{\omega}_g\}_{g \in E_n}) \right) \right].$$

(III) There is a $C > 0$ such that for all $n \in \mathbb{N}$

$$\text{Var} \left(\mathbf{W}_n(\{1 + \widehat{\omega}_g\}_{g \in E_n}) - \mathbf{W}_n(\{\mathbf{W}_{N-n}(\{1 + \widehat{\omega}_a\}_{a \triangleleft g})\}_{g \in E_n}) \right) \leq Cx^2(b/s)^n.$$

Hence the third term on the right side of (4.8) is bounded by $\epsilon\sqrt{C} \sup_{x \in \mathbb{R}} |F'(x)|$.

(I) For $g \in E_n$, define

$$Y_g^{(N)} := \overline{\mathbf{W}}_{N-n}(\{X_a^{(N)}\}_{a \triangleleft g}) \quad \text{and} \quad Z_g^{(N)} := \mathbf{W}_{N-n}(\{X_a^{(N)}\}_{a \triangleleft g}) - \overline{\mathbf{W}}_{N-n}(\{X_a^{(N)}\}_{a \triangleleft g}).$$

Notice that

$$\text{Var}(Y_g^{(N)}) = x_N(b/s)^n \quad \text{and} \quad \text{Var}(Z_g^{(N)}) \leq cx^2(b/s)^{2n},$$

where the inequality holds for some $c > 0$ and all N, n by part (ii) of Lemma 4.11. The variance of the difference between $\mathbf{W}_N(\{X_a^{(N)}\}_{a \in E_N})$ and $\mathbf{W}_n(\{\overline{\mathbf{W}}_{N-n}(\{X_g^{(N)}\}_{a \triangleleft g})\}_{g \in E_n})$ can be written as

$$\text{Var} \left(\mathbf{W}_n(\{Y_g^{(N)} + Z_g^{(N)}\}_{g \in E_n}) - \mathbf{W}_n(\{Y_g^{(N)}\}_{g \in E_n}) \right),$$

which is bounded by a constant multiple of $x^2(b/s)^n$ for all N and n by the bound in the proof of part (i) of Lemma 4.11.

(II): For $a \in E_N$ define $\mathbf{x}_a^{(N)} := (s/b)^{N/2} (X_a^{(N)} - 1)$. Then, by our assumptions on $X_a^{(N)}$, the random variables $\mathbf{x}_a^{(N)}$ are i.i.d. with mean zero, variance $(s/b)^N x_N$, and satisfy a Lindeberg condition. We can write $\overline{\mathbf{W}}_{N-n}(\{X_a^{(N)}\}_{a \triangleleft g})$ in the form

$$\overline{\mathbf{W}}_{N-n}(\{X_a^{(N)}\}_{a \triangleleft g}) = 1 + \underbrace{\left(\frac{b}{s}\right)^{\frac{n}{2}} \left[\frac{1}{(bs)^{\frac{N-n}{2}}} \sum_{a \triangleleft g} \mathbf{x}_a^{(N)} \right]}_{\text{(Central limit-type sum)}},$$

where the sum in square brackets includes $|g| = (sb)^{N-n}$ elements. By the central limit theorem, there is convergence in law as $N \rightarrow \infty$

$$\overline{\mathbf{W}}_{N-n}(\{X_a^{(N)}\}_{a \triangleleft g}) - 1 \xrightarrow{\mathcal{L}} \widehat{\omega}_g$$

since $\widehat{\omega}_g$ has distribution $\mathcal{N}(0, x(b/s)^n)$. The result then follows since the copies of $\overline{\mathbf{W}}_{N-n}(\{X_a^{(N)}\}_{a \triangleleft g})$ for different $g \in E_n$ are independent and $h(\mathbf{y}) := \mathbb{E}[F(\mathbf{W}_n(\{\mathbf{y}_g\}_{g \in E_n}))]$ is a continuous function of arrays $\mathbf{y} \in \mathbb{R}^{(bs)^n}$ (recall that $|E_n| = b^n s^n$).

(III): The argument is similar to (I). Define

$$Y_g := 1 + \widehat{\omega}_g \quad \text{and} \quad Z_g^{(N)} := \mathbf{W}_{N-n}(\{1 + \widehat{\omega}_a\}_{a \triangleleft g}) - \overline{\mathbf{W}}_{N-n}(\{1 + \widehat{\omega}_a\}_{a \triangleleft g}).$$

Notice that for $N > n$

$$Y_g = \overline{\mathbf{W}}_{N-n}(\{1 + \widehat{\omega}_a\}_{a \triangleleft g}), \quad \text{Var}(Y_g) = x(b/s)^n, \quad \text{and} \quad \text{Var}(Z_g^{(N)}) \leq cx^2(b/s)^{2n},$$

where the inequality holds for some $c > 0$ and all N, n by part (ii) of Lemma 4.11. Thus we can apply the inequalities in the proof of part (i) of Lemma 4.11 to bound (III) since

$$\begin{aligned} \text{Var}\left(\mathbf{W}_n(\{1 + \widehat{\omega}_g\}_{g \in E_n}) - \mathbf{W}_n(\{\mathbf{W}_{N-n}(\{1 + \widehat{\omega}_a\}_{a \triangleleft g})\}_{g \in E_n})\right) \\ = \text{Var}\left(\mathbf{W}_n(\{Y_g + Z_g^{(N)}\}_{g \in E_n}) - \mathbf{W}_n(\{Y_g\}_{g \in E_n})\right). \end{aligned}$$

□

4.2 Reducing the vertex problem to the edge problem

Now we work towards a proof of Theorem 2.1 by approximating the original partition function by the partition function for a model in which the randomness is placed on the edges and is thus applicable to Theorem 4.6.

Definition 4.12. Define $M_n : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ by

$$M_n(x) := \frac{1}{b} \left[(1+x)^s e^{(s-1)[\lambda(2\beta_n) - 2\lambda(\beta_n)]} - 1 \right].$$

Remark 4.13. Note that for $g \in G_{n-k,n}$ we have $\text{Var}(W_n(\beta_n; g)) = M_n^k(0)$.

The proof of the following technical lemma is in Section 8.1

Lemma 4.14. Let $\beta_n = \widehat{\beta}(b/s)^{n/2}$ and $g \in G_{m,n}$. There is a $C > 0$ such that for all $n \in \mathbb{N}$ and $m \leq n/2$

$$i) \left| \text{Var}(W_n(\beta_n; g)) - \widehat{\beta}^2 \frac{s-1}{s-b} (1 - (b/s)^m) (b/s)^{n-m} \right| \leq C(b/s)^n$$

$$ii) \mathbb{E} \left[|W_n(\beta_n; g) - 1|^4 \right] \leq C(b/s)^{2(n-m)}.$$

Theorem 4.15 (Variance convergence). Define $\alpha_n := \widehat{\beta}^2 \frac{s-1}{s-b} (b/s)^{\lfloor n/2 \rfloor}$. As $n \rightarrow \infty$

$$|M_n^n(0) - \widehat{M}^{\lceil \frac{n}{2} \rceil}(\alpha_n)| \rightarrow 0.$$

In particular, by combining the above with part (i) of Lemma 4.10, we have that $M_n^n(0)$ converges to $\mathfrak{v}_{s,b}(\widehat{\beta}^2 \frac{s-1}{s-b})$ with large n .

Proof. For $\gamma_n := M_n^{\lfloor n/2 \rfloor}(0)$, we have the relations

$$\begin{aligned} |M_n^n(0) - \widehat{M}^{\lceil \frac{n}{2} \rceil}(\alpha_n)| &= |M_n^{\lceil \frac{n}{2} \rceil}(\gamma_n) - \widehat{M}^{\lceil \frac{n}{2} \rceil}(\alpha_n)| \\ &\leq |M_n^{\lceil \frac{n}{2} \rceil}(\gamma_n) - \widehat{M}^{\lceil \frac{n}{2} \rceil}(\gamma_n)| + |\widehat{M}^{\lceil \frac{n}{2} \rceil}(\gamma_n) - \widehat{M}^{\lceil \frac{n}{2} \rceil}(\alpha_n)|. \end{aligned} \quad (4.9)$$

Thus we just need to bound the two terms on right side above. For the second term above,

$$|\widehat{M}^{\lceil \frac{n}{2} \rceil}(\gamma_n) - \widehat{M}^{\lceil \frac{n}{2} \rceil}(\alpha_n)| \leq |\gamma_n - \alpha_n| \frac{d}{dy} \left[\widehat{M}^{\lceil \frac{n}{2} \rceil}(y) \right] \Big|_{y=\max(\alpha_n, \gamma_n)} \leq C(b/s)^n,$$

where we have applied part (i) of Lemma 4.14 to control the difference $\gamma_n - \alpha_n$ and part (iii) of Lemma 4.9 to control the derivative of $\widehat{M}^{\lceil n/2 \rceil}$.

Now we will bound the first term on the right side of (4.9). For x in a bounded interval $[0, L]$, there is a $c > 0$ such that

$$\widehat{M}(x) \leq M_n(x) \leq \widehat{M}(x) + c(b/s)^n. \quad (4.10)$$

By using a telescoping sum, we can write the difference between $M_n^m(\gamma_n)$ and $\widehat{M}^m(\gamma_n)$ as

$$\begin{aligned} M_n^m(\gamma_n) - \widehat{M}^m(\gamma_n) &= \sum_{k=0}^{m-1} \left(\widehat{M}^{m-k-1}(M_n^{k+1}(\gamma_n)) - \widehat{M}^{m-k}(M_n^k(\gamma_n)) \right) \\ &\leq \sum_{k=0}^{m-1} \left[\frac{d}{dy} \widehat{M}^{m-k-1}(y) \Big|_{y=M_n^{k+1}(\gamma_n)} \right] \left(M_n(M_n^k(\gamma_n)) - \widehat{M}(M_n^k(\gamma_n)) \right). \end{aligned}$$

Let $\tilde{n} \in \mathbb{N}$ be the smallest m such that $M_n^m(\gamma_n) > \widehat{M}^{m+1}(\gamma_n)$ or $\widehat{M}^m(\gamma_n) > L$. For $m < \tilde{n}$ the above is

$$\begin{aligned} &\leq cm \left(\frac{b}{s} \right)^n \max_{0 \leq k < m} \frac{d}{dy} \widehat{M}^{m-k-1}(y) \Big|_{y=\widehat{M}^{k+2}(\gamma_n)} \\ &\leq cm \left(\frac{b}{s} \right)^n \max_{0 \leq k < m} \left(\frac{s}{b} \right)^{m-k} \exp \left\{ \frac{s-1}{1-\frac{b}{s}} \widehat{M}^m(\gamma_n) \right\}, \end{aligned}$$

where we have applied (iv) of Lemma 4.9 to bound the derivative of \widehat{M} . For $c' := c \exp \left\{ \frac{s-1}{1-\frac{b}{s}} L \right\}$, the last expression is

$$\leq c' m \left(\frac{b}{s} \right)^{\lfloor \frac{n}{2} \rfloor}.$$

Thus we have that

$$M_n^m(\gamma_n) \leq \widehat{M}^m(\gamma_n) + c' m (b/s)^{\lfloor \frac{n}{2} \rfloor}.$$

for $m < \tilde{n}$. The value \tilde{n} must be greater than n for $n \gg 1$ since $\widehat{M}^{m+1}(\gamma_n)$ has the lower bound

$$\widehat{M}^{m+1}(\gamma_n) = \widehat{M}(\widehat{M}^m(\gamma_n)) > \frac{s}{b} \widehat{M}^m(\gamma_n) > \widehat{M}^m(\gamma_n) + \frac{s-1}{b} \gamma_n, \quad (4.11)$$

and γ_n has order $(b/s)^{\lfloor n/2 \rfloor}$ by Lemma 4.14. Therefore the first term on the right side of (4.9) decays with order $n(b/s)^{\lfloor n/2 \rfloor}$ for large n . \square

It is convenient to define a model for which the disorder at “low” generation vertices is neglected. We define “lower” generation to be $\leq \lfloor n/2 \rfloor$. This effectively means that we can work on a problem of size $\lfloor n/2 \rfloor$ in which randomness is placed on edges.

Proof of Theorem 2.1. Recall that $\beta_n = \widehat{\beta}(b/s)^{n/2}$. For $g \in G_{\lfloor n/2 \rfloor, n} \equiv E_{\lfloor n/2 \rfloor}$, define $X_g^{(n)} := W_n(\beta_n; g)$. Notice that the array of random variables $\{X_g^{(n)}\}_{g \in E_{\lfloor n/2 \rfloor}}$ satisfies the conditions of Theorem 4.6 as a consequence of Lemma 4.14. We can write $W_n(\beta_n) = W_n(\beta_n; D_n)$ as

$$W_n(\beta_n) = \mathbf{W}_{\lfloor \frac{n}{2} \rfloor}(\{X_g^{(n)}\}_{g \in E_{\lfloor \frac{n}{2} \rfloor}}) + \left(W_n(\beta_n; D_n) - \mathbf{W}_{\lfloor \frac{n}{2} \rfloor}(\{X_g^{(n)}\}_{g \in E_{\lfloor \frac{n}{2} \rfloor}}) \right).$$

We will show below that the difference between $W_n(\beta_n; D_n)$ and $\mathbf{W}_{\lfloor n/2 \rfloor}(\{X_g^{(n)}\}_{g \in E_{\lfloor n/2 \rfloor}})$ converges in probability to zero. Thus the convergence in law of $W_n(\beta_n)$ to $L_{\widehat{\beta}^2 \frac{s-1}{s-b}}$ follows from Theorem 4.6.

The random variables $W_n(\beta_n; D_n) - \mathbf{W}_{\lfloor n/2 \rfloor}(\{X_g^{(n)}\}_{g \in E_{\lfloor n/2 \rfloor}})$ and $\mathbf{W}_{\lfloor n/2 \rfloor}(\{X_g^{(n)}\}_{g \in E_{\lfloor n/2 \rfloor}})$ are uncorrelated, so

$$\begin{aligned} \mathbb{E} \left[\left(W_n(\beta_n; D_n) - \mathbf{W}_{\lfloor \frac{n}{2} \rfloor}(\{X_g^{(n)}\}_{g \in E_{\lfloor \frac{n}{2} \rfloor}}) \right)^2 \right] &= \text{Var} \left(W_n(\beta_n; D_n) \right) - \text{Var} \left(\mathbf{W}_{\lfloor \frac{n}{2} \rfloor}(\{X_g^{(n)}\}_{g \in E_{\lfloor \frac{n}{2} \rfloor}}) \right) \\ &= M_n^n(0) - \widehat{M}^{\lfloor \frac{n}{2} \rfloor}(\alpha_n). \end{aligned}$$

The above converges to zero by Theorem 4.15. \square

4.3 Transition to strong disorder in the distributions $L_r^{b,s}$

Let $(L_r^{b,s})_{r \geq 0}$ be the family of distributions satisfying the properties (I)-(III) listed in Theorem 2.1. It is clear from the definition that $L_r^{b,s}$ converges to the delta mass at 1 as $r \searrow 0$. The limiting behavior as $r \rightarrow \infty$ is less obvious, although it is reasonable to expect that the distributions transition to strong disorder behavior, in other terms, $L_r^{b,s}$ converges to the delta mass at 0. This is indeed the case as the next lemma shows. Our proof will follow the homogeneous environment tilting method used by Lacoïn and Moreno in Section 5.2 of [27], applied to random weights on bonds rather than sites.

Lemma 4.16. $L_r^{b,s}$ converges weakly to a δ -distribution at zero as $r \rightarrow \infty$.

Proof. Let W_r have distribution $L_r^{b,s}$. It suffices to show that the fractional moment $\mathbb{E}[W_r^{1/2}]$ converges to zero as $r \rightarrow \infty$. By the proof of Lemma 4.1, W_r can be constructed through a weak limit $N \rightarrow \infty$ of

$$\mathbf{W}_N \left(\left\{ 1 + \sqrt{r} \left(\frac{b}{s} \right)^{\frac{N}{2}} \widehat{\omega}_a \right\}_{a \in E_N} \right), \quad (4.12)$$

where \mathbf{W}_N is defined as in Definition 4.3 and $\{\widehat{\omega}_g\}_{g \in E_N}$ is an array of independent random variables with distribution $\mathcal{N}(0, 1)$. The proof of Lemma 4.1 also holds with the random variables

$$W_r^{(N)} := \mathbf{W}_N \left(\left\{ \exp \left\{ \sqrt{r} \left(\frac{b}{s} \right)^{\frac{N}{2}} \widehat{\omega}_a - \frac{r}{2} \left(\frac{b}{s} \right)^N \right\} \right\}_{a \in E_N} \right)$$

in place of (4.12) without any serious changes, and we will do our computation with this expression. Allow us to abuse notation by identifying measures with their corresponding expectation symbols. Define a new measure $\tilde{\mathbb{E}}$ on the system with density

$$\frac{d\tilde{\mathbb{E}}}{d\mathbb{E}} := \exp \left\{ - \sum_{a \in E_N} \left((bs)^{-\frac{N}{2}} \hat{\omega}_a + \frac{1}{2} (bs)^{-N} \right) \right\}.$$

By Cauchy-Schwarz, we can write

$$\mathbb{E}[W_{r,N}^{\frac{1}{2}}] = \tilde{\mathbb{E}} \left[\frac{d\mathbb{E}}{d\tilde{\mathbb{E}}} W_{r,N}^{\frac{1}{2}} \right] \leq \tilde{\mathbb{E}} \left[\left(\frac{d\mathbb{E}}{d\tilde{\mathbb{E}}} \right)^2 \right]^{\frac{1}{2}} \tilde{\mathbb{E}}[W_{r,N}]^{\frac{1}{2}}. \quad (4.13)$$

However, computations with Gaussian integrals yield that

$$\tilde{\mathbb{E}} \left[\left(\frac{d\mathbb{E}}{d\tilde{\mathbb{E}}} \right)^2 \right]^{\frac{1}{2}} = e^{\frac{1}{2}} \quad \text{and} \quad \tilde{\mathbb{E}}[W_{r,N}]^{\frac{1}{2}} = e^{-\frac{1}{2}\sqrt{r}}.$$

The first equation above uses that $|E_N| = (bs)^N$ and the second uses that each path in Γ_N contains s^N edges. By applying these inequalities in (4.13), we get that $\mathbb{E}[W_{r,N}^{1/2}]$ is smaller than $\exp\{(1 - \sqrt{r})/2\}$. For any fixed $r \in \mathbb{R}^+$,

$$\mathbb{E}[W_{r,N}^{\frac{1}{2}}] \longrightarrow \mathbb{E}[W_r^{\frac{1}{2}}]$$

as $N \rightarrow \infty$ since $W_{r,N} \xrightarrow{L} W_r$ and the second moment of $W_{r,N}$ is uniformly bounded for all $N > 0$. Therefore, $\mathbb{E}[W_{r,N}^{1/2}]$ is smaller than $\exp\{(1 - \sqrt{r})/2\}$, which converges to zero as $r \rightarrow \infty$. \square

5 Results for the $b = s$ case

In this section we prove Theorem 2.5.

5.1 Definitions relevant for the $s = b$ case

Remark 5.1. When considering $\beta = \hat{\beta}/n \ll 1$, it will be convenient to define

$$R_n(\hat{\beta}; g) := \sqrt{n} \left(W_n \left(\frac{\hat{\beta}}{n}; g \right) - 1 \right).$$

An explicit recursive relation for $R_n(\hat{\beta}; g)$ can be deduced from (3.1) starting from $R_n(\hat{\beta}; g) = 0$ for $g \in G_{0,n}$:

$$R_n(\hat{\beta}; g) = \sqrt{n} \left[\frac{1}{b} \sum_{i=1}^b \prod_{1 \leq j \leq b} \left(1 + \frac{1}{\sqrt{n}} R_n(\hat{\beta}; g \times (i, j)) \right) \prod_{1 \leq j \leq b-1} E_{g \circ (i, j)} \left(\frac{\hat{\beta}}{n} \right) - 1 \right]. \quad (5.1)$$

Notice that for $n \gg 1$ the random variables satisfy the asymptotic relation

$$\begin{aligned} R_n(\hat{\beta}; g) &= \frac{1}{b} \sum_{i=1}^b \sum_{1 \leq j \leq b} R_n(\hat{\beta}; g \times (i, j)) + \frac{1}{b\sqrt{n}} \sum_{i=1}^b \sum_{1 \leq j_1 < j_2 \leq b} R_n(\hat{\beta}; g \times (i, j_1)) R_n(\hat{\beta}; g \times (i, j_2)) \\ &\quad + \frac{\hat{\beta}}{b\sqrt{n}} \sum_{i=1}^b \sum_{1 \leq j \leq b-1} \omega_{g \circ (i, j)} + (\text{Lower-order terms}). \end{aligned} \quad (5.2)$$

The above uses that $\omega_{g \diamond (i,j)}$ has mean zero through the approximation

$$E_{g \diamond (i,j)}\left(\frac{\hat{\beta}}{n}\right) = \frac{e^{\frac{\hat{\beta}}{n}\omega_{g \diamond (i,j)}}}{\mathbb{E}[e^{\frac{\hat{\beta}}{n}\omega_{g \diamond (i,j)}}]} \approx 1 + \frac{\hat{\beta}}{n}\omega_{g \diamond (i,j)} + O\left(\frac{1}{n^2}\right).$$

The third term on the right-hand side of (5.2) introduces independent noise at each step. The variance of $R_n(\hat{\beta}; g)$ is given by $\varrho_k^{(n)}(\hat{\beta}) = n\sigma_k(\hat{\beta}/n)$, where $\sigma_k(\beta)$ is defined as in Corollary 3.2.

Definition 5.2. We define $\hat{R}_n(\hat{\beta}; g)$ as the solution to the following recursive relation, which corresponds to (5.2) without the lower-order terms:

$$\begin{aligned} \hat{R}_n(\hat{\beta}; g) &:= \frac{1}{b} \sum_{i=1}^b \sum_{1 \leq j \leq b} \hat{R}_n(\hat{\beta}; g \times (i, j)) + \frac{1}{b\sqrt{n}} \sum_{i=1}^b \sum_{1 \leq j_1 < j_2 \leq b} \hat{R}_n(\hat{\beta}; g \times (i, j_1)) \hat{R}_n(\hat{\beta}; g \times (i, j_2)) \\ &\quad + \frac{\hat{\beta}}{b\sqrt{n}} \sum_{i=1}^b \sum_{1 \leq j \leq b-1} \omega_{g \diamond (i,j)}. \end{aligned} \quad (5.3)$$

The initial condition is $\hat{R}_n(\hat{\beta}; g) = 0$ for $g \in G_{0,n}$. In addition we define

$$\hat{R}_{k,n}(\hat{\beta}) := \frac{1}{b^{n-k}} \sum_{g \in G_{k,n}} \hat{R}_n(\hat{\beta}; g).$$

The lemma below follows from the recursive relation (5.3) that defines $\hat{R}_n(\hat{\beta}; g)$. We define

$$\hat{\varrho}_k^{(n)}(\hat{\beta}) := \mathbb{E}[\hat{R}_n(\hat{\beta}; g)^2].$$

for $g \in G_{k,n}$. Note that the expectation on the right is independent of g by the iid assumption.

Lemma 5.3. *The moments $\hat{\varrho}_k^{(n)}$ satisfy the recursive relation*

$$\hat{\varrho}_{k+1}^{(n)}(\hat{\beta}) = \hat{\varrho}_k^{(n)}(\hat{\beta}) + \frac{1}{n} \frac{b-1}{2b} [\hat{\varrho}_k^{(n)}(\hat{\beta})]^2 + \frac{1}{n} \frac{\hat{\beta}^2(b-1)}{b}.$$

The following lemma is also an immediate consequence of the recursive relation (5.3).

Lemma 5.4. *The random variables $\hat{R}_{k,n}(\hat{\beta})$ satisfy the recursive equation in $k \in \mathbb{N}$ given by:*

$$\begin{aligned} \hat{R}_{k+1,n}(\hat{\beta}) &= \hat{R}_{k,n}(\hat{\beta}) + \frac{1}{b\sqrt{n}} \sum_{g \in G_{k,n}} \sum_{i=1}^b \sum_{1 \leq j_1 < j_2 \leq b} \hat{R}_n(\hat{\beta}; g \times (i, j_1)) \hat{R}_n(\hat{\beta}; g \times (i, j_2)) \\ &\quad + \sum_{g \in G_{k,n}} \frac{\hat{\beta}}{b\sqrt{n}} \sum_{i=1}^b \sum_{1 \leq j \leq b-1} \omega_{g \diamond (i,j)}. \end{aligned}$$

5.2 The case of $b = s$ and $\hat{\beta} < \kappa_b$

We will now reset the notation for the map M_n that was assigned in Section 4 to take in consideration the new scaling $\beta_n = \hat{\beta}/n$ that is relevant when $b = s$.

Definition 5.5. Define the maps $M_n, \widehat{M}_n : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ for $n \in \mathbb{N}$ as

- $M_n(x) := \frac{n}{b} \left[\left(1 + \frac{x}{n}\right)^b e^{(b-1)[\lambda(2\frac{\hat{\beta}}{n}) - 2\lambda(\frac{\hat{\beta}}{n})]} - 1 \right]$

- $\widehat{M}_n(x) := x + \frac{1}{n} \frac{b-1}{2} x^2 + \frac{1}{n} \frac{\widehat{\beta}^2(b-1)}{b}$

Remark 5.6. The maps M_n and \widehat{M}_n are defined so that

$$\varrho_{k+1}^{(n)}(\widehat{\beta}) = M_n(\varrho_k^{(n)}(\widehat{\beta})) \quad \text{and} \quad \widehat{\varrho}_{k+1}^{(n)}(\widehat{\beta}) = \widehat{M}_n(\widehat{\varrho}_k^{(n)}(\widehat{\beta})).$$

Thus the sequence $k \mapsto \varrho_k^{(n)}$ is determined by repeated compositions of the maps M_n .

The proof of Lemma 5.7 is calculus-based, and we have put it in Section 8.2

Lemma 5.7. *Let $\widehat{\beta} \in (0, \kappa_b)$. Define $\gamma(x, y) := \tan^{-1} \left(\frac{\sqrt{b}}{\widehat{\beta}\sqrt{2}} x \right) + \frac{\widehat{\beta}(b-1)}{\sqrt{2b}} y$, and for $\epsilon > 0$ let*

$$A_\epsilon = \left\{ (x, m, n) \in \mathbb{R}^+ \times \mathbb{N} \times \mathbb{N} : \gamma \left(x, \frac{m}{n} \right) < \frac{\pi}{2} - \epsilon \right\}.$$

Then for any $\epsilon > 0$ there is a constant $C > 0$ such that

$$\begin{aligned} i) \quad & \sup_{(x, m, n) \in A_\epsilon} \frac{d}{dx} \widehat{M}_n^m(x) \leq C, \\ ii) \quad & \sup_{(x, m, n) \in A_\epsilon} \left| \widehat{M}_n^m(x) - \frac{\widehat{\beta}\sqrt{2}}{\sqrt{b}} \tan \left(\tan^{-1} \left(\frac{\sqrt{b}}{\widehat{\beta}\sqrt{2}} x \right) + \widehat{\beta} \frac{b-1}{\sqrt{2b}} \frac{m}{n} \right) \right| \leq \frac{C}{n}, \\ iii) \quad & \sup_{(x, m, n) \in A_\epsilon} \left| M_n^m(x) - \widehat{M}_n^m(x) \right| \leq \frac{C}{n}. \end{aligned}$$

Remark 5.8. The function $g(r) := \frac{\widehat{\beta}\sqrt{2}}{\sqrt{b}} \tan \left(\tan^{-1} \left(\frac{\sqrt{b}}{\widehat{\beta}\sqrt{2}} x \right) + \widehat{\beta} \frac{b-1}{\sqrt{2b}} r \right)$ is the solution to the differential equation

$$\frac{d}{dr} g(r) = \frac{b-1}{2} g^2(r) + \frac{b-1}{b} \widehat{\beta}^2$$

with initial condition $g(0) = x$.

Lemma 5.9.

i) For any $\widehat{\beta} \in [0, \kappa_b)$ and $r \in [0, 1]$, there is convergence as $n \rightarrow \infty$ given by

$$\widehat{\varrho}_{[nr]}^{(n)}(\widehat{\beta}) \longrightarrow \frac{\widehat{\beta}\sqrt{2}}{\sqrt{b}} \tan \left(\widehat{\beta} \frac{b-1}{\sqrt{2b}} r \right).$$

ii) For $m \in \mathbb{N}$, define $\widehat{\varrho}_{k,m}^{(n)}(\widehat{\beta}) := \mathbb{E}[(\widehat{R}_n(\widehat{\beta}; g))^m]$. Then for any $\widehat{\beta} \in [0, \kappa_b)$ and $m \in \mathbb{N}$, there is a $C > 0$ such that for all $r \in [0, 1]$ and $n \in \mathbb{N}$

$$\widehat{\varrho}_{[nr], 2m}^{(n)}(\widehat{\beta}) \leq C \left(\widehat{\varrho}_{[nr], 2}^{(n)}(\widehat{\beta}) \right)^m$$

and, in particular, $\widehat{\varrho}_{[nr], 2m}^{(n)}(\widehat{\beta})$ is uniformly bounded over the stated range of variables.

iii) For $0 \leq j \leq k \leq n$, the random variables $\widehat{R}_{j,n}(\widehat{\beta})$ and $\widehat{R}_{k,n}(\widehat{\beta}) - \widehat{R}_{j,n}(\widehat{\beta})$ are uncorrelated.

Proof. Part (i): This is a consequence of part (ii) of Lemma 5.7 and the recursive relation

$$\widehat{\varrho}_{k+1}^{(n)}(\widehat{\beta}) = M_n\left(\widehat{\varrho}_k^{(n)}(\widehat{\beta})\right).$$

Part (ii): As a consequence of part (i), $\widehat{\varrho}_{k,2}^{(n)}(\widehat{\beta})$ is uniformly bounded for all $n > 0$ and $0 \leq k \leq nr$. For $m = 3$ it is straightforward to show that $\widehat{\varrho}_{k,3}^{(n)}$ satisfies the recursive relation

$$\begin{aligned} \widehat{\varrho}_{k+1,3}^{(n)}(\widehat{\beta}) &= \frac{1}{b}\widehat{\varrho}_{k,3}^{(n)}(\widehat{\beta}) + \frac{1}{\sqrt{n}}\frac{3(b-1)}{b}[\widehat{\varrho}_{k,2}^{(n)}(\widehat{\beta})]^2 + \frac{1}{n}\frac{b-1}{b}\widehat{\varrho}_{k,2}^{(n)}(\widehat{\beta})\widehat{\varrho}_{k,3}^{(n)}(\widehat{\beta}) + \frac{1}{n^{\frac{3}{2}}}\frac{(b-1)(b-2)}{b^2}[\widehat{\varrho}_{k,2}^{(n)}(\widehat{\beta})]^3 \\ &\quad + \frac{1}{n^{\frac{3}{2}}}\frac{b-1}{2b^2}[\widehat{\varrho}_{k,3}^{(n)}(\widehat{\beta})]^2 + \frac{1}{n^{\frac{3}{2}}}\frac{\widehat{\beta}^3(b-1)}{b^2}\mathbb{E}[\omega^3]. \end{aligned}$$

From this we see that the absolute value of $\widehat{\varrho}_{k,3}^{(n)}(\widehat{\beta})$ satisfies a recursive inequality of the form

$$\begin{aligned} |\widehat{\varrho}_{k+1,3}^{(n)}(\widehat{\beta})| &\leq \frac{1}{b}|\widehat{\varrho}_{k,3}^{(n)}(\widehat{\beta})| + \frac{3}{\sqrt{n}}\frac{b-1}{b}[\widehat{\varrho}_{k,2}^{(n)}(\widehat{\beta})]^2 + \frac{1}{n}\frac{b-1}{b}\widehat{\varrho}_{k,2}^{(n)}(\widehat{\beta})|\widehat{\varrho}_{k,3}^{(n)}(\widehat{\beta})| + \frac{1}{n^{\frac{3}{2}}}\frac{\widehat{\beta}^3(b-1)}{b^2}\mathbb{E}[\omega^3] \\ &\quad + \frac{1}{n^{\frac{3}{2}}}\frac{(b-1)(b-2)}{b^2}[\widehat{\varrho}_{k,2}^{(n)}(\widehat{\beta})]^3 + \frac{1}{n^{\frac{3}{2}}}\frac{b-1}{2b^2}[\widehat{\varrho}_{k,3}^{(n)}(\widehat{\beta})]^2 \end{aligned}$$

with $|\widehat{\varrho}_{0,3}^{(n)}(\widehat{\beta})| = 0$. Let $\widehat{k} \in \mathbb{N}$ be the first value such that $|\widehat{\varrho}_{\widehat{k},3}^{(n)}(\widehat{\beta})| > 1$. There exists an $\epsilon \in (0, 1)$ and a $c > 0$ such that for all $k \leq \min(\widehat{k} - 1, nr)$ the above is

$$\leq \epsilon|\widehat{\varrho}_{k,3}^{(n)}(\widehat{\beta})| + \frac{c}{\sqrt{n}}.$$

It follows that $\sup_{0 \leq k \leq \lfloor rn \rfloor} |\widehat{\varrho}_{k+1,3}^{(n)}(\widehat{\beta})|$ is $O(\frac{1}{\sqrt{n}})$ and, in particular, bounded for all $n \gg 1$.

Similar reasoning holds for $m = 4$. From the recursive equality for $\widehat{\varrho}_{k,4}^{(n)}$, which is straightforward to derive but lengthy to write out, we have an inequality for $\widehat{\varrho}_{k,4}^{(n)}(\widehat{\beta})$ of the form

$$\widehat{\varrho}_{k+1,4}^{(n)}(\widehat{\beta}) \leq \epsilon\widehat{\varrho}_{k,4}^{(n)}(\widehat{\beta}) + \frac{6(b^2-1)}{b^2}[\widehat{\varrho}_{k,2}^{(n)}(\widehat{\beta})]^2 + \frac{c}{n}$$

for some $\epsilon \in (0, 1)$ and $c > 0$. It follows that $|\widehat{\varrho}_{\lfloor rn \rfloor, 4}^{(n)}(\widehat{\beta})|$ is bounded by a multiple of $[\widehat{\varrho}_{\lfloor rn \rfloor, 2}^{(n)}(\widehat{\beta})]^2$. The reasoning can be extended inductively to arbitrary moments.

Part (iii): This follows from the recursive formula in Lemma 5.4. □

Theorem 5.10. For $\beta \in (0, \kappa_b)$ and $n \in \mathbb{N}$, define the stochastic process

$$Y_r^{(n)}(\widehat{\beta}) := \widehat{R}_{\lfloor rn \rfloor, n}(\widehat{\beta}), \quad r \in [0, 1].$$

As $n \rightarrow \infty$, $(Y_r^{(n)}(\widehat{\beta}))_{r \in [0, 1]}$ converges in law to a continuous Gaussian process $(\mathbf{y}_r(\widehat{\beta}))_{r \in [0, 1]}$ with independent, mean zero increments and

$$\mathbf{y}_r(\widehat{\beta}) \stackrel{d}{=} \mathcal{N}\left(0, \frac{\widehat{\beta}\sqrt{2}}{\sqrt{b}} \tan\left(\widehat{\beta}\frac{b-1}{\sqrt{2b}}r\right)\right).$$

Remark 5.11. The limit process $(\mathbf{y}_r(\hat{\beta}))_{r \in [0,1]}$ is simply a deterministic time-change of a standard Brownian motion $(\mathbf{B}_t)_{t \geq 0}$

$$\mathbf{y}_r(\hat{\beta}) \stackrel{d}{=} \mathbf{B}_{\tau_r},$$

where $\tau_r := \frac{\hat{\beta}\sqrt{2}}{\sqrt{b}} \tan\left(\hat{\beta}\frac{b-1}{\sqrt{2b}}r\right)$.

Proof of Theorem 5.10. First we will focus on finite-dimensional distributional convergence. For all $n \in \mathbb{N}$ and $0 \leq r \leq s \leq 1$,

$$\mathbb{E}[Y_r^{(n)}(\hat{\beta})] = 0 \quad \text{and} \quad \mathbb{E}\left[Y_r^{(n)}(\hat{\beta})\left(Y_s^{(n)}(\hat{\beta}) - Y_r^{(n)}(\hat{\beta})\right)\right] = 0, \quad (5.4)$$

where the second equality follows from part (iii) of Lemma 5.9. Thus finite-dimensional convergence is implied by the convergence of the variances $\mathbb{E}[|Y_r^{(n)}(\hat{\beta})|^2]$.

Next we show that $Y_r^{(n)}(\hat{\beta})$ converges to a normal distribution for $r \in [0,1)$. By definition of $\hat{R}_{k,n}(\hat{\beta})$, we can write $Y_r^{(n)}(\hat{\beta})$ as

$$Y_r^{(n)}(\hat{\beta}) = \hat{R}_{\lfloor rn \rfloor, n}(\hat{\beta}) = \underbrace{\frac{1}{b^{n-\lfloor rn \rfloor}} \sum_{g \in G_{\lfloor rn \rfloor, n}} \hat{R}_n(\hat{\beta}; g)}_{\text{Lindeberg-Feller central limit-type sum}}. \quad (5.5)$$

Lindeberg-Feller central limit-type sum

However, the random variables $\hat{R}_n(\hat{\beta}; g)$ are i.i.d. with mean zero and variance $\hat{\varrho}_{\lfloor nr \rfloor}^{(n)}(\hat{\beta})$. Recall that $\hat{\varrho}_{\lfloor nr \rfloor}^{(n)}(\hat{\beta})$ converges to $\frac{\hat{\beta}\sqrt{2}}{\sqrt{b}} \tan\left(\hat{\beta}\frac{b-1}{\sqrt{2b}}r\right)$ as $n \rightarrow \infty$ by Lemma 5.9. Since $|G_{\lfloor rn \rfloor, n}| = b^{2(n-\lfloor rn \rfloor)}$ and by the above observation, (5.5) converges to a normal distribution with mean zero and variance $\frac{\hat{\beta}\sqrt{2}}{\sqrt{b}} \tan\left(\hat{\beta}\frac{b-1}{\sqrt{2b}}r\right)$ provided that the following Lindeberg condition is satisfied: as $n \rightarrow \infty$ for any fixed $\epsilon > 0$

$$\frac{1}{b^{2n-2\lfloor rn \rfloor}} \left[\sum_{g \in G_{\lfloor rn \rfloor, n}} \mathbb{E}\left[|\hat{R}_n(\hat{\beta}; g)|^2 \chi\left(|\hat{R}_n(\hat{\beta}; g)| > \epsilon b^{n-\lfloor rn \rfloor}\right)\right] \right] \rightarrow 0.$$

However, by Chebyshev, the expression above is bounded by

$$\frac{1}{\epsilon^2 b^{4n-4\lfloor rn \rfloor}} \sum_{g \in G_{\lfloor rn \rfloor, n}} \mathbb{E}\left[|\hat{R}_n(\hat{\beta}; g)|^4\right] = \frac{1}{\epsilon^2 b^{2n-2\lfloor rn \rfloor}} \hat{\varrho}_{\lfloor nr \rfloor, 4}^{(n)}(\hat{\beta}) \leq \frac{C}{\epsilon^2 b^{2n-2\lfloor rn \rfloor}}, \quad (5.6)$$

where the second inequality holds for some $C > 0$ and all $n > 0$ by part (ii) of Lemma 5.9. The right side of (5.6) converges to zero, so the Lindeberg condition holds.

When $r = 1$, the sum (5.5) contains only one term and thus does not immediately fall within the purview of Lindeberg-Feller. Let $(u_n)_{n \geq 0}$ be a non-decreasing sequence of integers with $0 \leq u_n \leq n$, $u_n \rightarrow \infty$, and $\frac{u_n}{n} \rightarrow 0$. Define $\hat{Y}_1^{(n)}(\hat{\beta}) := \hat{R}_{n-u_n, n}(\hat{\beta})$. By writing

$$Y_1^{(n)}(\hat{\beta}) = \hat{Y}_1^{(n)}(\hat{\beta}) + \left(Y_1^{(n)}(\hat{\beta}) - \hat{Y}_1^{(n)}(\hat{\beta})\right), \quad (5.7)$$

it is sufficient for us to show that

I) $\hat{Y}_1^{(n)}(\hat{\beta})$ converges to the normal distribution $\mathcal{N}\left(0, \frac{\hat{\beta}\sqrt{2}}{\sqrt{b}} \tan\left(\hat{\beta}\frac{b-1}{\sqrt{2b}}\right)\right)$.

II) $Y_1^{(n)}(\hat{\beta}) - \hat{Y}_1^{(n)}(\hat{\beta})$ converges in probability to zero.

The proof of I) follows analogously to the case of $r < 1$ above with $\lfloor rn \rfloor$ replaced by $n - u_n$. For II), we observe that the expectation of the square of the right side of (5.7) is

$$\begin{aligned}\mathbb{E}\left[\left|Y_1^{(n)}(\hat{\beta}) - \hat{Y}_1^{(n)}(\hat{\beta})\right|^2\right] &= n\mathbb{E}\left[\left|\hat{R}_{n,n}(\hat{\beta}) - \hat{R}_{n-u_n,n}(\hat{\beta})\right|^2\right] \\ &= \hat{\varrho}_n^{(n)}(\hat{\beta}) - \hat{\varrho}_{n-u_n}^{(n)}(\hat{\beta}) \\ &= O\left(\frac{u_n}{n}\right).\end{aligned}$$

The second equality holds because $\hat{R}_{n,n}(\hat{\beta}) - \hat{R}_{n-u_n,n}(\hat{\beta})$ and $\hat{R}_{n-u_n,n}(\hat{\beta})$ are uncorrelated, and the order equality follows from part (ii) of Lemma 5.9 since $\hat{\varrho}_n^{(m)}(\hat{\beta}) = \hat{M}_n^m(0)$. \square

Lemma 5.12. *There is a $C > 0$ such that for all $n \geq 1$*

$$\mathbb{E}\left[\left|R_n(\hat{\beta}) - \hat{R}_n(\hat{\beta})\right|^2\right] \leq \frac{C}{n}.$$

Proof. The random variables $R_n(\hat{\beta}) - \hat{R}_n(\hat{\beta})$ and $\hat{R}_n(\hat{\beta})$ are uncorrelated. This can be seen, for instance, by defining a series of intermediary random variables $\mathbf{R}_n^{(\ell)}(\hat{\beta})$ between $R_n(\hat{\beta})$ and $\hat{R}_n(\hat{\beta})$ indexed by $\ell \in [0, n]$ such that $\mathbf{R}_n^{(\ell)}(\hat{\beta}) := \mathbf{R}_n^{(\ell)}(\hat{\beta}; g)$ for $g = D_n$, where the family of random variables $\{\mathbf{R}_n^{(\ell)}(\hat{\beta}; g)\}_{g \in G_{k,n}}$ satisfies the recursive relation (5.1) of $R_n(\hat{\beta}; g)$ for $g \in G_{k,n}$ with $k < \ell$ and the recursive relation (5.2) of $\hat{R}_n(\hat{\beta}; g)$ for $g \in G_{k,n}$ with $k \geq \ell$. By this construction, $\hat{R}_n(\hat{\beta}) = \mathbf{R}_n^{(0)}(\hat{\beta})$ and $R_n(\hat{\beta}) = \mathbf{R}_n^{(n)}(\hat{\beta})$ and the increments $\mathbf{R}_n^{(\ell+1)}(\hat{\beta}) - \mathbf{R}_n^{(\ell)}(\hat{\beta})$ are pairwise uncorrelated.

Since $R_n(\hat{\beta}) - \hat{R}_n(\hat{\beta})$ and $\hat{R}_n(\hat{\beta})$ are uncorrelated, we have the first equality below:

$$\begin{aligned}\mathbb{E}\left[\left(R_n(\hat{\beta}) - \hat{R}_n(\hat{\beta})\right)^2\right] &= \mathbb{E}[R_n^2(\hat{\beta})] - \mathbb{E}[\hat{R}_n^2(\hat{\beta})] \\ &:= \varrho_n(\hat{\beta}) - \hat{\varrho}_n(\hat{\beta}) \\ &= M_n^n(0) - \hat{M}_n^n(0) = O(1/n).\end{aligned}\tag{5.8}$$

The order equality follows from part (iii) of Lemma 5.7. \square

Proof of Theorem 2.5 in the $b = s$ and $\beta < \kappa_b$ case. We can write

$$R_n(\hat{\beta}) = \hat{R}_n(\hat{\beta}) + (R_n(\hat{\beta}) - \hat{R}_n(\hat{\beta})),\tag{5.9}$$

where the rightmost term converges to zero in probability as a consequence of Lemma 5.12. Moreover, $\hat{R}_n(\hat{\beta})$ converges in distribution to $\mathcal{N}(0, v_b(\beta))$ for $v_b(\beta) := \frac{\hat{\beta}\sqrt{2}}{\sqrt{b}} \tan(\hat{\beta}\frac{b-1}{\sqrt{2b}})$ by Theorem 5.10, which completes the proof. \square

5.3 The case of $b = s$ and $\hat{\beta} = \kappa_b$

We will need slightly more refined estimates to treat the critical point case, $\hat{\beta} = \kappa_b$, than was required for the analysis of $\hat{\beta} < \kappa_b$. Recall that we defined the family of random variables $\hat{R}_n(\hat{\beta}; g)$ through a recursive formula (5.3) that was essentially a quadratic approximation of the recursive formula (5.1) satisfied by the family of random variables $R_n(\hat{\beta}; g) := \sqrt{n}(W_n(\hat{\beta}; g) - 1)$. At the critical point, a third-order term in the recursive formula for $R_n(\hat{\beta}; g)$ becomes nonnegligible, making it convenient to formulate the random variables $\tilde{R}_n(\hat{\beta}; g)$ in the definition below that serve the same purpose as $\hat{R}_n(\hat{\beta}; g)$ last section.

Definition 5.13. For $g \in G_{k,n}$ we define $\tilde{R}_n(\hat{\beta}; g)$ to be the solution to the recursive relation

$$\begin{aligned} \tilde{R}_n(\hat{\beta}; g) &:= \frac{1}{b} \sum_{i=1}^b \sum_{1 \leq j \leq b} \tilde{R}_n(\hat{\beta}; g \times (i, j)) + \frac{1}{b\sqrt{n}} \sum_{i=1}^b \sum_{1 \leq j_1 < j_2 \leq b} \prod_{k=1,2} \tilde{R}_n(\hat{\beta}; g \times (i, j_k)) \\ &+ \frac{1}{bn} \sum_{i=1}^b \sum_{1 \leq j_1 < j_2 < j_3 \leq b} \prod_{k=1,2,3} \tilde{R}_n(\hat{\beta}; g \times (i, j_k)) + \frac{\hat{\beta}}{b\sqrt{n}} \sum_{i=1}^b \sum_{1 \leq j \leq b-1} \omega_{g \circ (i,j)} \end{aligned} \quad (5.10)$$

with initial condition is $\tilde{R}_n(\hat{\beta}; g) = 0$ for $g \in G_{0,n}$.

In addition, we define $\widetilde{W}_n(\hat{\beta}; g)$, $\widetilde{R}_{k,n}(\hat{\beta})$, $\widetilde{M}_n(x)$, $\widetilde{\varrho}_{k,p}^{(n)}(\hat{\beta})$ in analogy to $\widehat{W}_n(\hat{\beta}; g)$, $\widehat{R}_{k,n}(\hat{\beta})$, $\widehat{M}_n(x)$, $\widehat{\varrho}_{k,p}^{(n)}(\hat{\beta})$ from the last section.

Definition 5.14. Let $0 \leq k \leq n$ and $g \in G_{k,n}$.

- $\widetilde{W}_n(\hat{\beta}; g) := 1 + \frac{1}{\sqrt{n}} \tilde{R}_n(\hat{\beta}; g)$
- $\widetilde{R}_{k,n}(\hat{\beta}) := \frac{1}{b^{n-k}} \sum_{g \in G_{k,n}} \tilde{R}_n(\hat{\beta}; g)$
- $\widetilde{M}_n(x) := x + \frac{b-1}{2n} x^2 + \frac{(b-1)(b-2)}{6n^2} x^3 + \frac{\hat{\beta}^2(b-1)}{bn}$
- $\widetilde{\varrho}_{k,p}^{(n)}(\hat{\beta}) := \mathbb{E}[(\tilde{R}_n(\hat{\beta}; g))^p]$

When $p = 2$ we identify $\widetilde{\varrho}_{k,2}^{(n)}(\hat{\beta}) \equiv \widetilde{\varrho}_k^{(n)}(\hat{\beta})$.

Remark 5.15. In analogy with (5.6), the maps \widetilde{M}_n satisfy

$$\widetilde{\varrho}_{k+1}^{(n)}(\hat{\beta}) = \widetilde{M}_n(\widetilde{\varrho}_k^{(n)}(\hat{\beta})). \quad (5.11)$$

The random variables $\widetilde{R}_{k,n}(\hat{\beta})$ have uncorrelated increments in the index $k \in \mathbb{N}$ and satisfy $\widetilde{R}_{n,n}(\hat{\beta}) = \widetilde{R}_n(\hat{\beta}; D_n)$.

Since the variance of $\widetilde{W}_n(\kappa_b/n; D_n)$ is $\widetilde{M}_n^n(0)$, the following lemma requires an analysis of n -fold compositions of the map \widetilde{M} . The proof is in Section 8.2.

Lemma 5.16. *As $n \rightarrow \infty$ the variance of the random variable $\widetilde{W}_n(\kappa_b/n; D_n)$ has the convergence*

$$\log(n) \text{Var} \left(\widetilde{W}_n \left(\frac{\kappa_b}{n}; D_n \right) \right) \longrightarrow \frac{6}{b+1}.$$

The following lemma contains some analogous results to Lemma 5.7, and its proof is placed in Section 8.2. The main point is (iii), which implies that the variance, $\widetilde{M}_n^n(0)$, of $\widetilde{W}_n(\hat{\beta}/n)$ is close to the variance, $\widetilde{M}_n^n(0)$, of $\widetilde{W}_n(\hat{\beta}/n)$.

Lemma 5.17. *Parts (i) and (ii) below hold for small enough $\epsilon > 0$.*

i) *There are constants, $c, C > 0$ such that for all $n > 1$*

$$c \frac{n}{\log n} \leq \widetilde{M}_n^{[n+\epsilon \log n]}(0) \leq C \frac{n}{\log n}.$$

ii) Define $\ell_n := \lceil n + \epsilon \log n \rceil$. There is $C > 0$ such that for all $n > 1$ all $m \leq \ell_n$

$$\left| \frac{d}{dx} \widetilde{M}_n^{\ell_n - m}(x) \right|_{x=\widetilde{M}_n^m(0)} \leq C \left(\frac{n}{\log n} \right)^2 \frac{1}{1 + \left(\frac{\pi}{b-1} \widehat{M}_n^m(0) \right)^2}.$$

iii) As $n \rightarrow \infty$,

$$\frac{\log n}{n} |M_n^n(0) - \widetilde{M}_n^n(0)| = O\left(\frac{1}{\log n}\right).$$

Proof of Theorem 2.5 in the case $s = b$ and $\beta = \kappa_b$. The reasoning follows along the same basic lines as for the case of $\beta < \kappa_b$. First we show that $R_n(\kappa_b) = \sqrt{n}(W_n(\kappa_b/n) - 1)$ can be approximated by $\widetilde{R}_n(\kappa_b; D_n) = \sqrt{n}(\widetilde{W}_n(\kappa_b/n; D_n) - 1)$, on the scale of $\sqrt{\log n}/\sqrt{n}$. Since $R_n(\kappa_b) - \widetilde{R}_n(\kappa_b; D_n)$ and $\widetilde{R}_n(\kappa_b; D_n)$ are uncorrelated

$$\begin{aligned} \mathbb{E} \left[\left(\left(\frac{\log n}{n} \right)^{\frac{1}{2}} R_n(\kappa_b) - \left(\frac{\log n}{n} \right)^{\frac{1}{2}} \widetilde{R}_n(\kappa_b; D_n) \right)^2 \right] &= \frac{\log n}{n} \left[\text{Var}(R_n(\kappa_b)) - \text{Var}(\widetilde{R}_n(\kappa_b; D_n)) \right] \\ &= \frac{\log n}{n} (M_n^n(0) - \widetilde{M}_n^n(0)) \rightarrow 0, \end{aligned}$$

where the convergence follows from part (iii) of Lemma 5.17. Thus the difference

$$\left(\frac{\log n}{n} \right)^{\frac{1}{2}} R_n\left(\frac{\kappa_b}{n}\right) - \left(\frac{\log n}{n} \right)^{\frac{1}{2}} \widetilde{R}_n\left(\frac{\kappa_b}{n}; D_n\right)$$

converges to zero in probability.

It remains to be shown that

$$Y_n := \left(\frac{\log n}{n} \right)^{\frac{1}{2}} \widetilde{R}_n(\kappa_b; D_n) = \left(\frac{\log n}{n} \right)^{\frac{1}{2}} \widetilde{R}_{n,n}(\kappa_b)$$

converges in law to a Gaussian. The proof is similar to the proof of Theorem 5.10 so we only provide a sketch. For an increasing sequence $u_n \in \mathbb{N}$ with $u_n = o((\log n)/n)$, define $\widetilde{Y}_n := (\frac{\log n}{n})^{1/2} \widetilde{R}_{n-u_n, n}(\kappa_b)$. It suffices to show that

$$I) \quad Y_n - \widetilde{Y}_n \xrightarrow{\mathcal{P}} 0 \quad \text{and} \quad II) \quad \widetilde{Y}_n \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{6}{b+1}\right).$$

For (I) $Y_n - \widetilde{Y}_n$ and \widetilde{Y}_n are uncorrelated, and thus by similar reasoning as above we get the equality

$$\mathbb{E}[(Y_n - \widetilde{Y}_n)^2] = \frac{\log n}{n} (\widetilde{M}_n^n(0) - \widetilde{M}_n^{n-u_n}(0)).$$

Using that the derivative of \widetilde{M}_n is increasing, the above is smaller than

$$\begin{aligned} &\leq \frac{\log n}{n} \widetilde{M}_n^{u_n}(0) \frac{d}{dx} \widetilde{M}_n^{n-u_n} \Big|_{x=\widetilde{M}_n^{u_n}(0)} \\ &\leq C u_n \frac{n}{\log n} \rightarrow 0. \end{aligned} \tag{5.12}$$

The second inequality holds for some $C > 0$ because with large n

$$\widetilde{M}_n^{u_n}(0) = u_n \frac{\widehat{\beta}^2 b - 1}{n b} + O\left(\frac{u_n}{n^2}\right),$$

and the derivative of $\widetilde{M}_n^{n-u_n}$ is bounded by a constant multiple of $n^2/\log^2 n$ by part (ii) of Lemma 5.17. The last line of (5.12) converges to zero by our assumption on u_n .

For (II) recall that \tilde{Y}_n is a normalized sum of i.i.d. random variables

$$\tilde{Y}_n := \frac{1}{b^{u_n}} \sum_{g \in G_{n-u_n, n}} \left(\frac{\log n}{n} \right)^{\frac{1}{2}} \tilde{R}_n(\kappa_b; g).$$

The random variables $(\frac{\log n}{n})^{1/2} \tilde{R}_n(\kappa_b; g)$ have variance

$$\text{Var} \left(\left(\frac{\log n}{n} \right)^{1/2} \tilde{R}_n(\kappa_b; g) \right) = \text{Var}(\tilde{Y}_n) \approx \text{Var}(\tilde{Y}_n) \rightarrow \frac{6}{b+1},$$

where the approximation holds by (5.12) and the convergence is by Lemma 5.16. Thus the convergence in law of \tilde{Y}_n to $\mathcal{N}(0, \frac{6}{b+1})$ follows from the Lindeberg-Feller central limit theorem if the random variables $(\frac{\log n}{n})^{1/2} \tilde{R}_n(\kappa_b; g)$ satisfy a Lindeberg condition. Uniform bounds on the fourth moments can be obtained by an analogous argument to that for part (ii) of Lemma 5.9. \square

5.4 Variance explosion

We now focus on the variance explosion stated in Theorem 2.5 for the case $s = b$ with $\hat{\beta} > \kappa_b$. In fact, Lemma 5.19 provides a stronger result, which narrows down where the variance begins to blow-up in terms of the system size. Before moving to the proof of Lemma 5.19, we prove the following lemma, which is merely a translation of previous results.

Lemma 5.18. *The variance of $\sqrt{\log n}(W_{\lfloor n\kappa_b/\hat{\beta} \rfloor}(\hat{\beta}/n) - 1)$ converges to $\frac{6}{b+1}$.*

Proof. In the build-up (see Lemmas 5.16 and 5.17) to the proof of (2.7) in Theorem 2.5 we showed that the variance of $\sqrt{\log n}(W_n(\kappa_b/n) - 1)$ converges to $\frac{6}{b+1}$. We can simply apply this result by replacing the system size n by $\lfloor n\kappa/\hat{\beta} \rfloor$ and the inverse temperature κ_b/n by $\kappa_b/\lfloor n\kappa/\hat{\beta} \rfloor \approx \hat{\beta}/n$. \square

Lemma 5.19. *Suppose $s = b$ with $\hat{\beta} > \kappa_b$. Then as n goes to infinity the variance of $W_n(\hat{\beta}/n)$ tends to infinity. Moreover, there are constants $0 < c_\downarrow < c_\uparrow$ such that for $\ell_\downarrow(n) := \lfloor n\kappa_b/\hat{\beta} + c_\downarrow \log n \rfloor$ and $\ell_\uparrow(n) := \lfloor n\kappa_b/\hat{\beta} + c_\uparrow \log n \rfloor$*

$$\text{Var} \left(W_{\ell_\downarrow(n)}(\hat{\beta}/n) \right) \rightarrow 0 \quad \text{and} \quad \text{Var} \left(W_{\ell_\uparrow(n)}(\hat{\beta}/n) \right) \rightarrow \infty.$$

Remark 5.20. In Theorem 2.11 of [27], Lacoin and Moreno state that there exist $c, C > 0$ such that for all $\beta > 0$

$$\exp \left\{ -\frac{c}{\beta^2} \right\} \leq \lambda(\beta) - p(\beta) \leq \exp \left\{ -\frac{C}{\beta} \right\}, \quad (5.13)$$

where $p(\beta)$ is the quenched free energy

$$p(\beta) := \lim_{n \rightarrow \infty} \mathbb{E} [\log (Z_n(\beta))] .$$

By following their argument, which relies on [27, Proposition 4.3], we can use the convergence of $\text{Var}(W_{\ell_\downarrow(n)}(\hat{\beta}/n))$ to zero to get an upper bound of the form

$$\lambda(\beta) - p(\beta) \leq \exp \left\{ -\frac{\kappa_b \log b}{\beta} + C \log \beta \right\} \quad (5.14)$$

for some $C > 0$ and all $\beta > 0$. However, since the inequality (5.13) involves qualitatively different bounds from above and below, it is not clear to us that (5.14) is truly a step in the right direction.

Proof of Lemma 5.19. The variance of $W_n(\widehat{\beta}/n)$ is equal to $\frac{1}{n}\varrho_n^{(n)}(\widehat{\beta})$, so it is sufficient to show that $\varrho_n^{(n)}(\widehat{\beta})$ eventually grows at a faster than linear rate in n . Recall that $\varrho_m^{(n)}(\widehat{\beta}) = M_n^m(0)$ and $M_n(x) \geq \widehat{M}_n(x)$ for all $x \geq 0$, where M_n and \widehat{M}_n are defined as in Definition 5.5. As a consequence of Lemma 5.18, we can choose n large enough so that $\widehat{M}_n^k(0)$ with $k := \lfloor n\kappa_b/\widehat{\beta} \rfloor$ is larger than any $L > 0$. Pick some L greater than $\frac{16\widehat{\beta}}{(\widehat{\beta}-\kappa_b)(b-1)}$. Notice that we have the lower bound $\varrho_n^{(n)}(\widehat{\beta}) > \widehat{M}_n^{n-k}(L)$. The first inequality below holds for any $K > 0$ by the form of the maps \widehat{M}_n :

$$\widehat{M}_n^r(K) > K \left(1 + \frac{1}{n} \frac{b-1}{2} K \right)^r > K \exp \left\{ \frac{r}{n} \frac{b-1}{4} K \right\}. \quad (5.15)$$

The second inequality holds as long as $\frac{K(b-1)}{2n}$ is smaller than the solution $x > 1$ to the equation $1 + x = \exp\{x/2\}$. Observe that the function $g(r) := \widehat{M}_n^r(K)$ doubles before r reaches the value $\lceil \frac{4n}{(b-1)K} \rceil$. Define the sequence $r_j \in \mathbb{N}$ such that

$$r_j := \sum_{i=1}^j \left\lceil \frac{4n}{2^{i-1}(b-1)L} \right\rceil.$$

Through repeated use of (5.15), we have that

$$\widehat{M}_n^{r_j}(L) > L2^j \quad \text{and thus} \quad \widehat{M}_n^{n-k}(L) > L2^{|S_{L,n}|} \quad (5.16)$$

for the set $S_{L,n} := \{j \in \mathbb{N} \mid r_j < n - k\}$. However, the number of terms in $S_{L,n}$ will grow linearly with large n :

$$\frac{|S_{L,n}|}{n} \approx \frac{n - k - \frac{8n}{(b-1)L}}{n} > \frac{1}{2} \left(1 - \frac{\kappa_b}{\widehat{\beta}} \right),$$

where the second inequality uses that L has a lower bound of $\frac{16\widehat{\beta}}{(\widehat{\beta}-\kappa_b)(b-1)}$. It follows that $\widehat{M}_n^n(0)$ grows exponentially in n and thus the variance of $W_n(\widehat{\beta}/n)$ does also.

In fact the variance of $W_k(\widehat{\beta}/n)$ already blows up as k reaches the value $\ell_\uparrow(n) := \lfloor n\kappa_b/\widehat{\beta} + c_\uparrow \log n \rfloor$ for $c_\uparrow > 2 \log_2(e)$ since with large n

$$\frac{|S_{L,\ell_n}|}{c_\uparrow \log n} \longrightarrow 1.$$

Thus,

$$\text{Var}\left(W_{\ell(n)}(\widehat{\beta}/n)\right) = \frac{1}{n}\varrho_{\ell(n)}^{(n)}(\widehat{\beta}) > \frac{1}{n}\widehat{M}_n^{\ell(n)-k}(L) > \frac{L}{n}2^{|S_{L,\ell(n)}|} \gg \sqrt{n}.$$

Finally, the existence of c_\downarrow such that $\text{Var}\left(W_{\ell_\downarrow(n)}(\widehat{\beta}/n)\right)$ converges to zero for $\ell_\downarrow(n) := \lfloor n\kappa_b/\widehat{\beta} + c_\downarrow \log n \rfloor$ follows from Lemma 5.18 and part (i) of Lemma 5.17. For the application of Lemma 5.17, the system size “ n ” in the statement of the lemma is replaced by our value $\lfloor n\kappa_b/\widehat{\beta} \rfloor$. \square

6 The case of $s < b$

In this section we prove Theorem 2.8 in the case $s < b$. The analysis here is trivial, but it is mathematically instructive to see what happens differently between the $b = s$ and $b > s$ cases. Recall that we are generically scaling the inverse temperature $\beta \equiv \beta_n$ to zero as the size, n , of the system grows.

It will be convenient to reset some of the notations from previous sections. For $g \in G_{k,n}$ define

$$R_n(g) := \frac{1}{\beta_n} \left(W_n(\beta_n; g) - 1 \right),$$

which satisfies the recursive relation

$$R_n(g) = \frac{1}{\beta_n} \left[\frac{1}{b} \sum_{i=1}^b \prod_{1 \leq j \leq s} \left(1 + \beta_n R_n(g \times (i, j)) \right) \prod_{1 \leq j \leq s-1} E_{g \diamond (i, j)}(\beta_n) - 1 \right]. \quad (6.1)$$

Definition 6.1. For $g \in G_{k,n}$ we define $\widehat{R}_n(g)$ as the solution to the recursive relation

$$\widehat{R}_n(g) = \frac{1}{b} \sum_{i=1}^b \sum_{1 \leq j \leq s} \widehat{R}_n(g \times (i, j)) + \frac{1}{b} \sum_{i=1}^b \sum_{1 \leq j \leq s-1} \omega_{g \diamond (i, j)} \quad (6.2)$$

with $\widehat{R}_n(g) = 0$ for $g \in G_{0,n}$.

Remark 6.2. The recursive relation (6.2) does not contain any quadratic terms in the random variables $\widehat{R}_n(g \times (i, j))$ as it did in the last section. Notice that when $g = D_n$ the recursive relation (6.2) simply gives us

$$\widehat{R}_n(D_n) = \sum_{m=1}^n \frac{1}{b^m} \sum_{a \in V_m} \omega_a.$$

Definition 6.3. Define the maps $M_n, \widehat{M}_n : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ for $n \in \mathbb{N}$ as

- $M_n(x) := \frac{1}{b\beta_n^2} \left[(1 + x\beta_n^2)^s e^{(s-1)[\lambda(2\beta_n) - 2\lambda(\beta_n)]} - 1 \right]$
- $\widehat{M}_n(x) := \frac{s}{b}x + \frac{s-1}{b}$

Remark 6.4. As usual the maps M_n and \widehat{M}_n are defined so that

$$\mathbb{E} \left[(R_n(D_n))^2 \right] = M_n^n(0) \quad \text{and} \quad \mathbb{E} \left[(\widehat{R}_n(D_n))^2 \right] = \widehat{M}_n^n(0).$$

The following lemma has an easy proof, which we do not include.

Lemma 6.5. *There is a $C > 0$ such that the inequality below holds for all $n \geq 1$*

$$|M_n^n(0) - \widehat{M}_n^n(0)| \leq C\beta_n.$$

Proof of Theorem 2.8 in the case of $s < b$. By the recursive relation (6.2), it is clear that

$$\widehat{R}_n(D_n) \xrightarrow{\mathcal{P}} \sum_{m=1}^{\infty} \frac{1}{b^m} \sum_{a \in V_m} \omega_a.$$

Thus it is sufficient to show that the variance of the difference $R_n(D_n) - \widehat{R}_n(D_n)$ converges to zero with large n . The difference $R_n(D_n) - \widehat{R}_n(D_n)$ is uncorrelated with $\widehat{R}_n(D_n)$, so

$$\begin{aligned} \mathbb{E} \left[(R_n(D_n) - \widehat{R}_n(D_n))^2 \right] &= \mathbb{E} \left[(R_n(D_n))^2 \right] - \mathbb{E} \left[(\widehat{R}_n(D_n))^2 \right] \\ &= M_n^n(0) - \widehat{M}_n^n(0) = O\left(\frac{1}{n}\right), \end{aligned} \quad (6.3)$$

where the order equality holds by Lemma 6.5.

□

7 The edge model

A closely related polymer model can be defined by placing random i.i.d. weights ω_a on the edges $a \in E_n$ rather than the vertices of the diamond lattice D_n . The normalized partition function takes the form

$$W_n(\beta) := \frac{1}{|\Gamma_n|} \sum_{p \in \Gamma_n} \prod_{a \triangleleft p} E(\beta; a) \quad (7.1)$$

for $E(\beta; a) := \frac{e^{\beta\omega_a}}{\mathbb{E}[e^{\beta\omega_a}]}$, where $a \triangleleft p$ means that the edge a lies along the path p . In analogy with (2.2) there is an obvious recursive relation between the distributions of $W_n(\beta)$ and $W_{n+1}(\beta)$:

$$W_{n+1}(\beta) \stackrel{d}{=} \frac{1}{b} \sum_{i=1}^b \prod_{1 \leq j \leq s} W_n^{(i,j)}(\beta),$$

where the $W_n^{(i,j)}(\beta)$ are independent copies of $W_n(\beta)$. In the case of $b = s$, there is an interesting dichotomy between the edge and vertex models that is revealed by considering the expected number of edges/vertices shared between two randomly chosen paths: two random paths picked from Γ_n have an expected number of shared edges equal to 1 and shared vertices equal to $\frac{b-1}{b}n$. This is a consequence of the hierarchical positions of the vertices on the graph in which the lower generation vertices are shared by exponentially more paths. Indeed, the edges play an equivalent role with respect to paths to only the n^{th} generation vertices. Since the elements $\prod_{a \triangleleft p} E(\beta; a)$ in the sum (7.1) share fewer of the variables $E(\beta; a)$ on average in the edge model, this suggests that the edge model will exhibit “less disorder” and perhaps even have a finite range of small β in which the model is weakly disordered. On the other hand the second moment of $W_n(\beta)$ tends to infinity with large n since it satisfies the recursive equation

$$\mathbb{E}[W_{n+1}^2(\beta)] = \frac{1}{b} \mathbb{E}[W_n^2(\beta)] + \frac{b-1}{b}.$$

with $\mathbb{E}[W_0^2(\beta)] = \mathbb{E}[E^2(\beta; a)] > 1$. The following theorem follows in analogy to Theorem 2.5.

Theorem 7.1 ($b = s$, edge model). *With large $n \in \mathbb{N}$ the variance of $W_n(\hat{\beta}/\sqrt{n})$ converges to zero for $\hat{\beta} \leq \kappa_b$ and tends to infinity for $\hat{\beta} > \kappa_b$. When $\hat{\beta} < \kappa_b$ we have the following weak convergence:*

$$\sqrt{n} \left(W_n \left(\frac{\hat{\beta}}{\sqrt{n}} \right) - 1 \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, v_b(\hat{\beta})),$$

where the variance of the limit is $v_b(\hat{\beta}) := (1/\hat{\beta}^2 - 1/\kappa_b^2)^{-1}$. At the critical value $\hat{\beta} = \kappa_b$, the limit result becomes:

$$\sqrt{\log n} \left(W_n \left(\frac{\hat{\kappa}_b}{\sqrt{n}} \right) - 1 \right) \xrightarrow{\mathcal{L}} \mathcal{N} \left(0, \frac{6}{b+1} \right).$$

8 Miscellaneous proofs

8.1 Proofs from Section 4

Proof of Lemma 4.9. Note that there is a $c > 0$ such that for all $x \in \mathbb{R}^+$

$$\frac{s}{b}x < \widehat{M}(x) \leq \frac{s}{b}xe^{cx}. \quad (8.1)$$

It is sufficient to show that the inequalities (i)-(iii) hold for $\lambda > 0$ small enough. We will assume that λ satisfies

$$\exp \left\{ \frac{c \frac{s}{b} \lambda}{1 - \frac{b}{s}} \right\} \leq \frac{s}{b}. \quad (8.2)$$

Part (i): Let $\tilde{n} \equiv \tilde{n}(x, N) \in \mathbb{N}$ be the first value of m such that

$$\widehat{M}^m(x(b/s)^N) > x \frac{s}{b}.$$

When N is large, the value of \tilde{n} must be greater than N by the following induction argument: For $m < \tilde{n} \wedge N$, we can apply (8.1) to get the first two inequalities below.

$$\begin{aligned} \widehat{M}^{m+1}(x(b/s)^N) &\leq x(b/s)^{N-m-1} \exp \left\{ c \sum_{r=0}^m \widehat{M}^r(x(b/s)^N) \right\} \\ &\leq x(b/s)^{N-m-1} \exp \left\{ c \sum_{r=0}^m \left(\frac{b}{s} \right)^{m-r} \widehat{M}^m(x(b/s)^N) \right\} \\ &= x(b/s)^{N-m-1} \exp \left\{ \frac{c}{1 - \frac{b}{s}} \widehat{M}^m(x(b/s)^N) \right\} \\ &\leq x(b/s)^{N-m-1} \exp \left\{ \frac{c \frac{s}{b} x}{1 - \frac{b}{s}} \right\} \leq x \left(\frac{b}{s} \right)^{N-m}. \end{aligned} \quad (8.3)$$

The third inequality uses the assumption that $m < \tilde{n}$, and the last inequality uses that $x \leq \lambda$. The above implies that $N + 1 \leq \tilde{n} \wedge N$. We can apply analogous inequalities as above to get the desired bound for $\widehat{M}^{N-n}(x(b/s)^N)$.

Part (ii): By similar reasoning as in (8.3),

$$\widehat{M}^{N-n}(x(b/s)^N) - x(b/s)^n \leq x(b/s)^n \left(\exp \left\{ \frac{c}{1 - \frac{b}{s}} \widehat{M}^{N-n}(x(b/s)^N) \right\} - 1 \right).$$

Moreover, applying the inequality $e^r - 1 \leq r e^r$ yields

$$\begin{aligned} &\leq x(b/s)^n \frac{c}{1 - \frac{b}{s}} \widehat{M}^{N-n}(x(b/s)^N) \exp \left\{ \frac{c}{1 - \frac{b}{s}} \widehat{M}^{N-n}(x(b/s)^N) \right\} \\ &\leq C x^2 (b/s)^{2n}, \end{aligned}$$

where the last inequality holds for some $C > 0$ as a consequence of (i).

Parts (iii) and (iv): Let $y := x(b/s)^n$. To bound the derivative, we can apply the chain rule to get

$$\begin{aligned} \frac{d}{dx} \widehat{M}^n(x(b/s)^n) &\leq \frac{s}{b} \prod_{r=0}^{n-1} \left(1 + \widehat{M}^r(y) \right)^{s-1} \\ &= \frac{s}{b} \prod_{r=1}^n \left(1 + \widehat{M}^{r-n}(\widehat{M}^{n-1}(y)) \right)^{s-1} \\ &\leq \frac{s}{b} \exp \left\{ (s-1) \sum_{r=1}^n \widehat{M}^{r-n}(\widehat{M}^{n-1}(y)) \right\}. \end{aligned}$$

Moreover, since $\widehat{M}^{-1}(y) \leq \frac{b}{s}y$,

$$\leq \frac{s}{b} \exp \left\{ \frac{s-1}{1-\frac{b}{s}} \widehat{M}^{n-1}(y) \right\}.$$

Note that the above gives us (iv). We can apply that $\widehat{M}^{n-1}(x(b/s)^n) \leq x(s/b)$, which holds by the reasoning in the proof of (i), to bound the above by

$$\leq \frac{s}{b} \exp \left\{ x \frac{s}{b} \frac{s-1}{1-\frac{b}{s}} \right\} \leq \frac{s}{b} \exp \left\{ \lambda \frac{s}{b} \frac{s-1}{1-\frac{b}{s}} \right\}. \quad (8.4)$$

The above gives us a bound independent of n . □

Proof of Lemma 4.14. Part (i): Recall that $\text{Var}(W_n(\beta_n; g)) = M_n^m(0)$. Pick $c > 0$ to be large. For $x < c(b/s)^{n/2}$, there is a $c' > 0$ such that

$$0 \leq M_n(x) - \frac{s}{b}x - \widehat{\beta}^2 \frac{s-1}{b} \left(\frac{b}{s}\right)^n \leq c' \left(\frac{b}{s}\right)^{\frac{3n}{2}} + c' x \left(\frac{b}{s}\right)^{\frac{n}{2}}. \quad (8.5)$$

By using a telescoping sum, we can write

$$M_n^m(0) - \widehat{\beta}^2 \frac{s-1}{b} \frac{1 - (\frac{b}{s})^m}{1 - \frac{b}{s}} \left(\frac{b}{s}\right)^{n-m+1} = \sum_{k=0}^{m-1} \left(\frac{b}{s}\right)^{m-k-1} \left(M_n^{k+1}(0) - \frac{s}{b} M_n^k(0) - \widehat{\beta}^2 \frac{s-1}{b} \left(\frac{b}{s}\right)^n \right).$$

For any m such that $M_n^{m-1}(0) < c(b/s)^{n/2}$, (8.5) implies that the above is bounded by

$$\begin{aligned} &\leq c' \sum_{k=0}^{m-1} \left(\frac{b}{s}\right)^{m-k-1} \left(\left(\frac{b}{s}\right)^{\frac{3n}{2}} + \left(\frac{b}{s}\right)^{\frac{n}{2}} M_n^k(0) \right) \\ &\leq \frac{c'}{1-\frac{b}{s}} \left(\frac{b}{s}\right)^{\frac{3n}{2}} + \frac{c}{1-\frac{b}{s}} \left(\frac{b}{s}\right)^{\frac{n}{2}} M_n^m(0), \end{aligned}$$

where in the last inequality we have used that $M_n^k(0)$ is increasing with k . It follows that $M_n^m(0)$ is very close to $\widehat{\beta}^2 \frac{s-1}{s-b} (1 - (b/s)^m) (b/s)^{n-m}$ for all $m \leq n/2$ and in particular (i) holds.

Part (ii): Define $\sigma_{k,n}^{(m)} := \mathbb{E}[|W_n(\beta_n; g) - 1|^m]$. The recursive relation (3.1) and part (i) implies that there is a $C > 0$ such that for $k < m$ $|\sigma_{k,3}^{(n)}|$ is bounded by a constant multiple of $(b/s)^{2n}$. Thus, by applying (3.1) again, there is a $C > 0$ such that $\sigma_{k,4}^{(n)}$ satisfies

$$\sigma_{k+1,4}^{(n)} \leq \frac{s}{b^3} \sigma_{k,4}^{(n)} + \frac{6s(s-1)}{b^3} (\sigma_{k,2}^{(n)})^2 + C \left(\frac{b}{s}\right)^{2n}.$$

By applying part (i) we get the inequality. □

8.2 Proofs from Section 5

We remind the reader that the notations M_n and \widehat{M}_n refer to different maps in Section 5 than in Section 4.

Proof of Lemma. 5.7. Part (i): Define $r_m \in [0, \kappa_b)$ as

$$r_m := \frac{\sqrt{2b}}{\hat{\beta}(b-1)} \tan^{-1} \left(\frac{\sqrt{b}}{\hat{\beta}\sqrt{2}} \widehat{M}_n^m(x) \right). \quad (8.6)$$

Notice that the recursive form for $\widehat{M}_n^m(x)$ implies

$$\tan \left(\hat{\beta} \frac{b-1}{\sqrt{2b}} r_{m+1} \right) = \tan \left(\hat{\beta} \frac{b-1}{\sqrt{2b}} r_m \right) + \frac{1}{n} \hat{\beta} \frac{b-1}{\sqrt{2b}} \sec^2 \left(\hat{\beta} \frac{b-1}{\sqrt{2b}} r_m \right).$$

The above resembles a linear approximation around r_m with $r_{m+1} \approx r_m + \frac{1}{n}$. Since the derivative of tangent is increasing over the interval $[0, \pi)$, we will always have $r_{m+1} < r_m + \frac{1}{n}$. It follows that

$$\widehat{M}_n^m(x) \leq \frac{\hat{\beta}\sqrt{2}}{\sqrt{b}} \tan \left(\tan^{-1} \left(\frac{\sqrt{b}}{\hat{\beta}\sqrt{2}} x \right) + \hat{\beta} \frac{b-1}{\sqrt{2b}} \frac{m}{n} \right). \quad (8.7)$$

Since $\widehat{M}_n'(x) = 1 + \frac{b-1}{n}x$, by the chain rule we have

$$\begin{aligned} \frac{d}{dx} [\widehat{M}_n^m(x)] &= \prod_{j=1}^m \widehat{M}_n'(\widehat{M}_n^{j-1}(x)) = \prod_{j=1}^m \left(1 + \frac{b-1}{n} \widehat{M}_n^{j-1}(x) \right) \\ &\leq \exp \left\{ \frac{b-1}{n} \sum_{j=1}^m \widehat{M}_n^{j-1}(x) \right\} \\ &\leq \exp \left\{ \frac{\hat{\beta}\sqrt{2}(b-1)}{\sqrt{b}} \int_0^{\frac{m}{n}} \tan \left(\tan^{-1} \left(\frac{\sqrt{b}}{\hat{\beta}\sqrt{2}} x \right) + \hat{\beta} \frac{b-1}{\sqrt{2b}} r \right) dr \right\}. \end{aligned} \quad (8.8)$$

The second inequality above holds by (8.7) and a Riemann upper bound. The right side of (8.8) is bounded uniformly for all $x \geq 0$ and $m, n \in \mathbb{N}$ with $\gamma(x, \frac{m}{n}) < \pi/2 - \epsilon$.

Part (ii): Define the function

$$\phi(x, y) = \frac{\hat{\beta}\sqrt{2}}{\sqrt{b}} \tan \left(\tan^{-1} \left(\frac{\sqrt{b}}{\hat{\beta}\sqrt{2}} x \right) + \hat{\beta} \frac{b-1}{\sqrt{2b}} y \right).$$

The difference between $\widehat{M}_n^m(x)$ and $\phi(x, \frac{m}{n})$ can be bounded through a telescoping sum:

$$\begin{aligned} \left| \phi\left(x, \frac{m}{n}\right) - \widehat{M}_n^m(x) \right| &\leq \sum_{k=0}^{m-1} \left| \widehat{M}_n^{m-k-1} \left(\phi\left(x, \frac{k+1}{n}\right) \right) - \widehat{M}_n^{m-k} \left(\phi\left(x, \frac{k}{n}\right) \right) \right| \\ &\leq \left[\sup_{0 \leq \ell \leq m} \frac{d}{da} \widehat{M}_n^{m-\ell-1} \left(\phi\left(x, \frac{\ell}{n}\right) \right) \right] \sum_{k=0}^{m-1} \left| \phi\left(x, \frac{k+1}{n}\right) - \phi\left(x, \frac{k}{n}\right) \right| \\ &\leq C \sum_{k=0}^{m-1} \left| \int_{\frac{k}{n}}^{\frac{k+1}{n}} dr \left(r - \frac{k}{n} \right) \frac{\partial^2}{\partial r^2} \phi(x, r) \right| \\ &\leq \frac{C}{n} \int_0^{\frac{m}{n}} dr \frac{\partial^2}{\partial r^2} \phi(x, r) = \frac{C}{n} \frac{\partial}{\partial r} \phi(x, r) \Big|_{r=\frac{m}{n}} \leq \frac{C}{n} \frac{\partial}{\partial y} \phi(x, y) \leq \frac{\widehat{C}}{n}. \end{aligned} \quad (8.9)$$

In the third inequality above, we have used part (i) to bound the supremum since

$$\gamma\left(\phi\left(x, \frac{\ell}{n}\right), \frac{m-\ell-1}{n}\right) = \gamma\left(x, \frac{m-1}{n}\right) < \frac{\pi}{2} - \epsilon, \quad 0 \leq \ell \leq m.$$

The third inequality also uses the observation that

$$\widehat{M}_n\left(\phi\left(x, \frac{k}{n}\right)\right) = \phi\left(x, \frac{k}{n}\right) + \frac{1}{n} \frac{d}{dr} \phi(x, r) \Big|_{r=\frac{k}{n}}$$

and a second-order Taylor expansion to write

$$\phi\left(x, \frac{k+1}{n}\right) = \widehat{M}_n\left(\phi\left(x, \frac{k}{n}\right)\right) + \int_{\frac{k}{n}}^{\frac{k+1}{n}} dr \left(r - \frac{k}{n}\right) \frac{\partial^2}{\partial r^2} \phi(x, r).$$

The final inequality in (8.9) holds for some \widehat{C} by another application of part (i).

Part (iii): We begin with the observation that for any bounded interval $z \in [0, a]$ there is a $c > 0$ such that for all $n \geq 1$

$$\left| M_n(z) - \widehat{M}_n(z) \right| \leq \frac{c}{n^2}. \quad (8.10)$$

We can write the difference between $M_n^m(x)$ and $\widehat{M}_n^m(x)$ using the following telescoping sum:

$$\begin{aligned} M_n^m(x) - \widehat{M}_n^m(x) &= \sum_{k=1}^m \left[\widehat{M}_n^{m-k}(M_n^k(x)) - \widehat{M}_n^{m-k+1}(M_n^{k-1}(x)) \right] \\ &\leq \sum_{k=1}^m \left[\sup_{0 \leq z \leq M_n^k(x)} \frac{d}{dz} \left[\widehat{M}_n^{m-k}(z) \right] \right] \left[M_n(M_n^{k-1}(x)) - \widehat{M}_n(M_n^{k-1}(x)) \right]. \end{aligned} \quad (8.11)$$

Let $N \in \mathbb{N}$ be the smallest number such that

$$\max_{0 \leq \ell \leq N} \gamma\left(M_n^\ell(x), \frac{N - \ell}{n}\right) \geq \frac{\pi}{2} - \frac{\epsilon}{2}. \quad (8.12)$$

The above implies that $M_n^k(x)$ is bounded by some $c > 0$ for all $k < N$ and $n > 0$. Combining (8.11) with the remark (8.10) yields that there is a $C > 0$ such that for $\ell < N$

$$M_n^\ell(x) - \widehat{M}_n^\ell(x) \leq \frac{c}{n^2} \sum_{k=1}^{\ell} \left[\sup_{0 \leq z \leq M_n^k(x)} \frac{d}{dz} \left[\widehat{M}_n^{\ell-k}(z) \right] \right] \leq \frac{c' \ell}{n^2},$$

where the second inequality holds by part (ii). Plugging the above back into (8.12) with N replaced by m yields

$$\begin{aligned} \max_{0 \leq \ell \leq m} \gamma\left(M_n^\ell(x), \frac{m - \ell}{n}\right) &\leq \max_{0 \leq \ell \leq m} \gamma\left(\widehat{M}_n^\ell(x) + \frac{c'}{n}, \frac{m - \ell}{n}\right) \\ &\leq \max_{0 \leq \ell \leq m} \gamma\left(\widehat{M}_n^\ell(x), \frac{m - \ell}{n}\right) + \frac{c' \sqrt{b} \ell}{\widehat{\beta} \sqrt{2} n^2}. \end{aligned}$$

By our previous observation that $\widehat{M}_n^\ell(x) \leq \phi\left(x, \frac{\ell}{n}\right)$, the above is smaller than

$$\begin{aligned} &\leq \max_{0 \leq \ell \leq m} \gamma\left(\phi\left(x, \frac{\ell}{n}\right), \frac{m - \ell}{n}\right) + \frac{c' \sqrt{b} m}{\widehat{\beta} \sqrt{2} n^2} \\ &\leq \frac{\pi}{2} - \epsilon + \frac{c' \sqrt{b} 1}{\widehat{\beta} \sqrt{2} n} < \frac{\pi}{2} - \frac{\epsilon}{2}, \end{aligned}$$

where we have used the identity $\gamma(\phi(x, \frac{k}{n}), \frac{m-k}{n}) = \gamma(x, \frac{m}{n})$ and our assumption that $\gamma(x, \frac{m}{n}) \leq \pi/2 - \epsilon$. The last inequality holds for large enough n . It follows that $m < N$ and

$$M_n^\ell(x) - \widehat{M}_n^\ell(x) \leq \frac{\mathbf{c}'m}{n^2} = O\left(\frac{1}{n}\right).$$

□

Proof of Lemma 5.17. For notational convenience, we will equate \widetilde{M}_n^r for $r \in \mathbb{R}^+$ with $\widetilde{M}_n^{\lceil r \rceil}$ in this proof.

Part (i): From the proof of Lem 5.16 we know that for $n \gg 1$

$$\widetilde{M}_n^n(0) \approx \frac{6}{b+1} \frac{n}{\log n} < \frac{12}{b+1} \frac{n}{\log n},$$

where the inequality holds for n large enough. Consider the sequence of numbers $x_m^{(n)} := \widetilde{M}_n^m(\widetilde{M}_n^n(0))$ and $c_m^{(n)} := x_m^{(n)} \frac{\log n}{n}$. Notice that $c_{m+1}^{(n)} \frac{n}{\log n} = x_{m+1}^{(n)} = \widetilde{M}_n(x_m^{(n)})$, and so

$$\begin{aligned} c_{m+1}^{(n)} \frac{n}{\log n} &= c_m^{(n)} \frac{n}{\log n} \left(1 + \frac{(b-1)(c_m^{(n)})^2}{2 \log n} + \frac{(b-1)(b-2)(c_m^{(n)})^4}{6 \log^2 n} \right) + \frac{b-1}{bn} \\ &< c_m^{(n)} \frac{n}{\log n} \left(1 + \frac{(b-1)(c_m^{(n)})^2}{\log n} + \frac{(b-1)(b-2)(c_m^{(n)})^4}{6 \log^2 n} \right) \\ &< c_m^{(n)} \frac{n}{\log n} \exp \left\{ \frac{(b-1)(c_m^{(n)})^2}{\log n} + \frac{(b-1)(b-2)(c_m^{(n)})^4}{6 \log^2 n} \right\}. \end{aligned} \quad (8.13)$$

For all $m \in \mathbb{N}$ such that $c_m^{(n)} < U$ for some constant U , then $c_m^{(n)} \leq \frac{12}{b+1} \exp \left\{ m \frac{(b-1)U^2}{\log n} + \frac{(b-1)(b-2)U^4}{6 \log^2 n} \right\}$. Thus for $\epsilon > 0$ small enough so that $U > \frac{6}{b+1} \exp \{ \epsilon(b-1)U^2 \}$, then $x_m^{(n)} = \widetilde{M}_n^{n+m}(0)$ will be bounded by a constant multiple $\frac{n}{\log n}$ for all $m < \epsilon \log n$.

Getting a lower bound is easier since we can use the first line of (8.13) to conclude that

$$c_m^{(n)} > \left(1 + \frac{(b-1)(c_0^{(n)})^2}{2 \log n} \right)^m \approx \exp \left\{ \epsilon \frac{b-1}{2} (c_0^{(n)})^2 \right\},$$

where the approximation is for $m = \epsilon \log n$ and n large.

Part (ii): Define $r_m \in [0, 1]$ as

$$r_m := \frac{2}{\pi} \tan^{-1} \left(\frac{b-1}{\pi} \widetilde{M}_n^m(x) \right). \quad (8.14)$$

The chain rule gives the first equality below:

$$\frac{d}{dx} \widetilde{M}_n^{\ell_n-m}(x) \Big|_{x=\widetilde{M}_n^m(0)} = \prod_{j=1}^{\ell_n-m} \widetilde{M}_n' \left(\widetilde{M}_n^{m+j-1}(0) \right).$$

Since $\widetilde{M}_n'(x) = 1 + \frac{b-1}{n}x + \frac{(b-1)(b-2)}{2n^2}x^2$, the above is equal to

$$\begin{aligned} &= \prod_{j=1}^{\ell_n-m} \left(1 + \frac{b-1}{n} \widetilde{M}_n^{m+j-1}(0) + \frac{(b-1)(b-2)}{2n^2} \left(\widetilde{M}_n^{m+j-1}(0) \right)^2 \right) \\ &\leq \exp \left\{ \frac{b-1}{n} \left(1 + \frac{c}{\log n} \right) \sum_{j=1}^{\ell_n-m} \widetilde{M}_n^{m+j-1}(0) \right\}, \end{aligned}$$

where we have applied part (i) to bound $\widetilde{M}_n^{m+j-1}(0)$ by $\frac{cn}{\log n}$ in the quadratic term. By definition of the values $r_j \in [0, 1)$, we have the equality

$$= \exp \left\{ \frac{\pi}{n} \left(1 + \frac{c}{\log n} \right) \sum_{j=m}^{\ell_n-1} \tan \left(\frac{\pi}{2} r_j \right) \right\}. \quad (8.15)$$

The difference between $\frac{1}{n} \sum_{j=m}^{\ell_n-1} \tan \left(\frac{\pi}{2} r_j \right)$ and $\int_{r_m}^{r_{\ell_n}} \tan \left(\frac{\pi}{2} s \right) ds$ is uniformly bounded, so we have $C > 0$ such that

$$\begin{aligned} &\leq C \exp \left\{ \pi \left(1 + \frac{c}{\log n} \right) \int_{r_m}^{r_{\ell_n}} \tan \left(\frac{\pi}{2} s \right) ds \right\} \\ &\leq C \exp \left\{ \pi \left(1 + \frac{c}{\log n} \right) \left(\log \left(\cos \left(\frac{\pi}{2} r_m \right) \right) - \log \left(\cos \left(\frac{\pi}{2} r_{\ell_n} \right) \right) \right) \right\} \\ &\leq C' \frac{n^2}{\log^2 n} \frac{1}{1 + \left(\frac{\pi}{b-1} \widetilde{M}_n^m(0) \right)^2}. \end{aligned} \quad (8.16)$$

The last inequality applies the relations

$$\frac{1}{\cos^2 \left(\frac{\pi}{2} r_{\ell_n} \right)} = 1 + \tan^2 \left(\frac{\pi}{2} r_{\ell_n} \right) = 1 + \frac{\pi^2}{(b-1)^2} \left(\widetilde{M}_n^{\ell_n}(0) \right)^2 = O \left(\frac{n^2}{\log^2 n} \right),$$

where the order equality is by part (i).

Part (iii): There is a $C > 0$ such that for $0 < x < n$

$$M_n(x) - \widetilde{M}_n(x) \leq \frac{Cx^2}{n^2} \left(\widetilde{M}_n(x) - x \right) + \frac{C}{n^2}.$$

By using a telescoping sum, as in Part (iii) of Lemma 5.7, we can bound the difference between $M_n^m(0)$ and $\widetilde{M}_n^m(0)$ through

$$\begin{aligned} &M_n^m(0) - \widetilde{M}_n^m(0) \\ &\leq \sum_{k=1}^m \left[\frac{d}{dz} \left[\widetilde{M}_n^{m-k}(z) \right] \Big|_{z=M_n^k(x)} \right] \left[M_n(M_n^{k-1}(0)) - \widetilde{M}_n(M_n^{k-1}(0)) \right] \\ &\leq \sum_{k=1}^m \left[\frac{d}{dz} \left[\widetilde{M}_n^{m-k}(z) \right] \Big|_{z=M_n^k(x)} \right] \left[\frac{C(M_n^{k-1}(0))^2}{n^2} \left(\widetilde{M}_n(M_n^{k-1}(0)) - M_n^{k-1}(0) \right) + \frac{C}{n^2} \right]. \end{aligned} \quad (8.17)$$

Let $\epsilon > 0$ be small and let $\alpha_n \in \mathbb{N}$ be the minimum of $n+1$ and the first $m \in \mathbb{N}$ such that

$$M_n^m(0) > \widetilde{M}_n^{m+\epsilon \log n}(0). \quad (8.18)$$

For $k \leq m < \alpha_n$, part (ii) implies the third inequality below

$$\begin{aligned} \frac{d}{dz} \left[\widetilde{M}_n^{m-k}(z) \right] \Big|_{z=M_n^k(0)} &\leq \frac{d}{dz} \left[\widetilde{M}_n^{m-k}(z) \right] \Big|_{z=\widetilde{M}_n^{k+\epsilon \log n}(0)} \leq \frac{d}{dz} \left[\widetilde{M}_n^{n-k}(z) \right] \Big|_{z=\widetilde{M}_n^{k+\epsilon \log n}(0)} \\ &\leq C' \frac{n^2}{\log^2 n} \frac{1}{1 + \left(\frac{\pi}{b-1} \widetilde{M}_n^{k+\epsilon \log n-1}(0) \right)^2}. \end{aligned} \quad (8.19)$$

The first inequality above uses that $M_n^m(0) \leq \widetilde{M}_n^{m+\epsilon \log n}(0)$ and the fact that the derivative of \widetilde{M}_n is an increasing function, and second inequality uses that $\widetilde{M}_n(x) > x$.

Moreover, for $k < \alpha_n$, we have the first inequality below:

$$\begin{aligned} (M_n^{k-1}(0))^2 \left(\widetilde{M}_n(M_n^{k-1}(0)) - M_n^{k-1}(0) \right) \\ \leq \left(\widetilde{M}_n^{k+\epsilon \log n-1}(0) \right)^2 \left(\widetilde{M}_n^{k+\epsilon \log n}(0) - \widetilde{M}_n^{k+\epsilon \log n-1}(0) \right). \end{aligned} \quad (8.20)$$

Combining (8.17), (8.19), and (8.20) for $m < \alpha_n$ we have that

$$\begin{aligned} M_n^m(0) - \widetilde{M}_n^m(0) &\leq \frac{CC'}{\log^2 n} \sum_{k=1}^m \frac{\left(\widetilde{M}_n^{k+\epsilon \log n-1}(0) \right)^2 \left(\widetilde{M}_n^{k+\epsilon \log n}(0) - \widetilde{M}_n^{k+\epsilon \log n-1}(0) \right) + c}{1 + \left(\frac{\pi}{b-1} \widetilde{M}_n^{k+\epsilon \log n-1}(0) \right)^2} \\ &\leq \frac{CC'}{\log^2 n} \sum_{k=1}^{m+\epsilon \log n} \left[\frac{b-1}{\pi} (\widetilde{M}_n^k(0) - \widetilde{M}_n^{k-1}(0)) + c \right] \\ &= \frac{CC'}{\log^2 n} \left[\frac{b-1}{\pi} \widetilde{M}_n^{m+\epsilon \log n}(0) + c(m + \epsilon \log n) \right] \leq \frac{C''n}{\log^2 n}, \end{aligned}$$

where the final inequality holds for some $C'' > 0$ by part (i) and the constraint $m < \alpha_n \leq n+1$.

Hence, we have shown that for $m < \alpha_n$

$$M_n^m(0) - \widetilde{M}_n^m(0) \leq \frac{C'''n}{\log^2 n}.$$

However, by part (i), there is a $c > 0$ such that for large n

$$\widetilde{M}_n^{n+\epsilon \log n}(0) > c \frac{n}{\log n}.$$

It follows that $\alpha_n = n+1$ for large enough n . Thus we have shown that

$$\frac{\log n}{n} (M_n^n(0) - \widetilde{M}_n^n(0)) = O\left(\frac{1}{\log(n)}\right).$$

□

Proof of Lemma 5.16. By the remark (5.11), $\widetilde{W}_m(\kappa_b/n; D_n) = 1 + \widetilde{R}_n(\kappa_b/n; D_m)$ is related to the maps $\widetilde{M}_n : [0, \infty) \rightarrow [0, \infty)$ as follows:

$$\text{Var} \left[\widetilde{W}_n\left(\frac{\kappa_b}{n}; D_n\right) \right] = \mathbb{E} \left[\left| \widetilde{R}_n\left(\frac{\kappa_b}{n}; D_n\right) \right|^2 \right] =: \widetilde{\varrho}_n^{(n)}(\widehat{\beta}) = \widetilde{M}_n^n(0).$$

Define the sequence of numbers $r_m^{(n)} \in [0, 1)$ such that

$$r_m^{(n)} := \frac{2}{\pi} \tan^{-1} \left(\frac{b-1}{\pi} \widetilde{M}_n^m(0) \right).$$

It is enough for us to prove that

$$1 - r_n^{(n)} = \frac{b+1}{3(b-1)} \frac{\log n}{n} + O\left(\frac{1}{n}\right)$$

since, in that case,

$$\widetilde{M}_n^n(0) = \frac{\pi}{b-1} \tan\left(\frac{\pi}{2} r_n^{(n)}\right) \approx \frac{\pi}{b-1} \frac{1}{\frac{\pi}{2} - \frac{\pi}{2} r_n^{(n)}} = \frac{6}{b+1} \frac{n}{\log n} + O\left(\frac{n}{\log^2 n}\right).$$

In the analysis below, we will abuse notation by writing r_m in place of $r_m^{(n)}$. As a consequence of (5.11), the r_m 's obey the recursive equation

$$r_{m+1} = \frac{2}{\pi} \tan^{-1} \left(\tan \left(\frac{\pi}{2} r_m \right) + \frac{\pi}{2n} \sec^2 \left(\frac{\pi}{2} r_m \right) + \frac{\pi^2(b-2)}{6(b-1)n^2} \tan^3 \left(\frac{\pi}{2} r_m \right) \right). \quad (8.21)$$

Moreover, a second-order application of Taylor's theorem to the function $f(x) = \tan(\frac{\pi}{2}x)$ at the point $x = r_m$ guarantees that there is a value r_m^* in the interval $(r_m, r_m + \frac{1}{n})$ such that

$$r_m + \frac{1}{n} = \frac{2}{\pi} \tan^{-1} \left(\tan \left(\frac{\pi}{2} r_m \right) + \frac{\pi}{2n} \sec^2 \left(\frac{\pi}{2} r_m \right) + \frac{\pi^2}{4n^2} \tan \left(\frac{\pi}{2} r_m^* \right) \sec^2 \left(\frac{\pi}{2} r_m^* \right) \right). \quad (8.22)$$

Define $\Delta_m \equiv \Delta_m^{(n)}$ as

$$\Delta_m := \frac{\pi^2}{4n^2} \tan \left(\frac{\pi}{2} r_m^* \right) \sec^2 \left(\frac{\pi}{2} r_m^* \right) - \frac{\pi^2(b-2)}{6(b-1)n^2} \tan^3 \left(\frac{\pi}{2} r_m \right).$$

The above can be written as follows:

$$= \frac{2}{\pi n^2} \frac{b+1}{3(b-1)} \frac{1}{(1-r_m)^3} + O\left(\frac{1}{n^3(1-r_m)^4}\right).$$

By a second-order application of Taylor's theorem to the function $g(x) = \frac{2}{\pi} \tan^{-1}(x)$ in (8.22) at the point $x = \tan(\frac{\pi}{2} r_{m+1})$, there is an $r_m^{**} \in (r_{m+1}, r_m + \frac{1}{n})$ such that

$$r_m + \frac{1}{n} = r_{m+1} + \frac{2}{\pi} \Delta_m \frac{1}{1 + \tan^2 \left(\frac{\pi}{2} r_{m+1} \right)} - \frac{2}{\pi} \Delta_m^2 \frac{\tan \left(\frac{\pi}{2} r_m^{**} \right)}{\left(1 + \tan^2 \left(\frac{\pi}{2} r_m^{**} \right) \right)^2}. \quad (8.23)$$

To bound the difference between 1 and r_n , we can apply (8.23) as follows:

$$\begin{aligned} 1 - r_n &= \sum_{m=0}^{n-1} \left(r_m + \frac{1}{n} - r_{m+1} \right) \\ &= \frac{2}{\pi} \sum_{m=0}^{n-1} \Delta_m \cos^2 \left(\frac{\pi}{2} r_{m+1} \right) - \frac{2}{\pi} \sum_{m=0}^{n-1} \Delta_m^2 \sin \left(\frac{\pi}{2} r_m^{**} \right) \cos^3 \left(\frac{\pi}{2} r_m^{**} \right) \\ &= \frac{1}{n^2} \frac{b+1}{3(b-1)} \sum_{m=0}^{n-1} \frac{1}{1-r_m} + O\left(\frac{1}{n^4} \sum_{m=0}^{n-1} \frac{1}{(1-r_m)^3}\right). \end{aligned}$$

The first term above is a Riemann sum and the second term is smaller by a factor of $\frac{1}{n}$.

$$\begin{aligned} &= \frac{1}{n} \frac{b+1}{3(b-1)} \left(1 + O\left(\frac{n^{-1}}{1-r_n}\right) \right) \int_0^{r_n} \frac{1}{1-x} dx + O\left(\frac{1}{n^3} \int_0^{r_n} \frac{1}{(1-x)^3} dx\right) \\ &= \frac{1}{n} \frac{b+1}{3(b-1)} \left(1 + O\left(\frac{n^{-1}}{1-r_n}\right) \right) \ln \left(\frac{1}{1-r_n} \right) \end{aligned}$$

The above implies that $1 - r_n = \frac{b+1}{3(b-1)} \frac{\log n}{n} + O\left(\frac{1}{n}\right)$, which completes the proof. \square

References

- [1] T. Alberts, K. Khanin, J. Quastel: *The intermediate disorder regime for directed polymers in dimension $1 + 1$* , Ann. Probab. **42**, No. 3, 1212-1256 (2014).
- [2] T. Alberts, K. Khanin, J. Quastel: *The continuum directed random polymer*, J. Stat. Phys. **154**, No. 1-2, 305-326 (2014).
- [3] T. Alberts, M. Ortgiese: *The near-critical scaling window for directed polymers on disordered trees*, Electron. J. Probab. **18**, No. 19, 1-24 (2013).
- [4] G. Amir, I. Corwin, and J. Quastel: *Probability distribution of the free energy of the continuum directed random polymer in $1+1$ dimensions*, Comm. Pure Appl. Math. **64**, No. 4, 466-537 (2011).
- [5] J.D. Biggins: *Martingale convergence in the branching random walk*, J. Appl. Probab. **14**, No. 1, 25-37 (1977).
- [6] E. Bolthausen: *A note on the diffusion of directed polymers in a random environment*, Comm. Math. Phys. **123** 529-534 (1989).
- [7] E. Buffet, A. Patrick, J.V. Pulé: *Directed polymers on trees: a martingale approach*, J. Phys. A **26**, No. 8, 1823-1834 (1993).
- [8] F. Caravenna, R. Sun, and N. Zygouras: *Polynomial chaos and scaling limits of disordered systems*, to appear in J. Eur. Math. Soc.
- [9] F. Caravenna, R. Sun, and N. Zygouras: *The continuum disordered pinning model*, to appear in Prob. Theor. Rel. Fields.
- [10] F. Comets, T. Shiga, and N. Yoshida: *Probabilistic analysis of directed polymers in a random environment*, Adv. Stud. Pure Math. **39**, 115-142 (2004).
- [11] F. Comets, and N. Yoshida: *Directed polymers in random environment are diffusive at weak disorder*, Ann. Probab. **34**, No. 5, 1746-1770 (2006).
- [12] J. Cook, B. Derrida: *Polymers on disordered hierarchical lattices: a nonlinear combination of random variables*, J. Stat. Phys. **57** 89-139 (1989).
- [13] I. Corwin, A. Hammond: *KPZ line ensemble*, arXiv:1312.2600 [math.PR] (2013).
- [14] B. Derrida, E. Gardner: *Renormalisation group study of a disordered model*, J. Phys. A: Math. Gen. **17**, 3223-3236 (1984).
- [15] B. Derrida, G. Giacomin, H. Lacoïn, F.L. Toninelli: *Fractional moment bounds and disorder relevance for pinning models*, Commun. Math. Phys. **287**, 867-887 (2009).
- [16] B. Derrida, R.B. Griffith: *Directed polymers on disordered hierarchical lattices*, Europhys. Lett. **8**, No. 2, 111-116 (1989).
- [17] B. Derrida, V. Hakim, J. Vannimenius: *Effect of disorder on two-dimensional wetting*, J. Stat. Phys. **66** 1189-1213 (1992).
- [18] G.R.M. Flores, T. Seppäläinen, B. Valkó: *Fluctuation exponents for directed polymers in the intermediate disorder regime*, Electron. J. Probab. **19**, no. 89., 1-28 (2014).

- [19] T. Garel, C. Monthus: *Critical points of quadratic renormalizations of random variables and phase transitions of disordered polymer models on diamond lattices*, Phys. Rev. E **77**, 021132 (2008).
- [20] G. Giacomin, H. Lacoin, F.L. Toninelli: *Hierarchical pinning models, quadratic maps, and quenched disorder*, Probab. Theor. Rel. Fields **145**, (2009).
- [21] R.B. Griffith, M. Kaufman: *Spin systems on hierarchical lattices. Introduction and thermodynamical limit*, Phys. Rev. B, **3** 26, no. 9, 5022-5032 (1982).
- [22] B.M. Hambly, J.H. Jordan: *A random hierarchical lattice: the series-parallel graph and its properties*, Adv. Appl. Prob., **36**, 824-838 (2004).
- [23] B.M. Hambly, T. Kumagai: *Diffusion on the scaling limit of the critical percolation cluster in the diamond hierarchical lattice*, Adv. Appl. Prob., **36**, 824-838 (2004).
- [24] Y. Hu, Z. Shi: *Minimal position and critical martingale convergence in branching random walks, and directed polymers on disordered trees*, Ann. Probab., **37**, No. 2, 742-789 (2009).
- [25] J. Z. Imbrie, T. Spencer: *Diffusion of directed polymers in a random environment*, J. Stat. Phys. **52**, 609-622 (2001).
- [26] J.-P. Kahane, J. Peyrière: *Sur certaines martingales de Benoit Mandelbrot*, Adv. in Math. **22**, No. 2, 131-145 (1976).
- [27] H. Lacoin, G. Moreno: *Directed Polymers on hierarchical lattices with site disorder*, Stoch. Proc. Appl. **120**, No. 4, 467-493 (2010).
- [28] H. Lacoin: *Hierarchical pinning model with site disorder: disorder is marginally relevant*, Probab. Theor. Rel. Fields **148**, No. 1-2, 159-175 (2010).
- [29] H. Lacoin: *New bounds for the free energy of directed polymers in dimension $1 + 1$ and $1 + 2$* , Commun. Math. Phys. **294**, No. 2, 471-503 (2010).
- [30] P. Mörters, M. Ortgiese: *Minimal supporting subtrees for the free energy of polymers on disordered trees*, J. Math. Phys. **49**, No. 12, 125203, 21 pp. (2008).
- [31] S. Roux, A. Hansen, L.R. Silva, L.S. Lucena, and R. Pandey: *Minimal path on the hierarchical diamond lattice*, Journal of Statistical Physics **65**, No.1, 183-204 (1991).
- [32] T. Seppäläinen: *Scaling for a one-dimensional directed polymer with boundary conditions*, Ann. Probab. **40**, No. 1, 19-73 (2012).